Central Lancashire Online Knowledge (CLoK)

Title	Effects of irrelevant speech on semantic and phonological judgments of Chinese characters
Type	Article
URL	https://clok.uclan.ac.uk/50889/
DOI	##doi##
Date	2024
Citation	Meng, Zhu, Yan, Guoli, Marsh, John Everett orcid iconORCID: 0000-0002-9494-1287 and Liversedge, Simon Paul orcid iconORCID: 0000-0002-8579-8546 (2024) Effects of irrelevant speech on semantic and phonological judgments of Chinese characters. Journal of Cognitive Psychology . ISSN 2044-5911
Creators	Meng, Zhu, Yan, Guoli, Marsh, John Everett and Liversedge, Simon Paul

It is advisable to refer to the publisher's version if you intend to cite from the work. ##doi##

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

Effects of irrelevant speech on semantic and phonological judgments of Chinese characters

Zhu Meng, Guoli Yan, John E. Marsh & Simon P. Liversedge

Published online: 28 Mar 2024.

Submit your article to this journal ⌕

View related articles ⌕

View Crossmark data ⌕

Routledge
Taylor & Francis Group

Effects of irrelevant speech on semantic and phonological judgments of Chinese characters

Zhu Meng [a,b], Guoli Yan[b], John E. Marsh [c,d] and Simon P. Liversedge [c]

aSchool of Education Science, Jiangsu Normal University, Xuzhou, People's Republic of China; bFaculty of Psychology, Tianjin Normal University, Tianjin, People's Republic of China; cSchool of Psychology and Humanities, University of Central Lancashire, Preston, UK; dDepartment of Health, Learning and Technology, Luleå University of Technology, Luleå, Sweden

ABSTRACT
This study investigated whether background speech impairs lexical processing and how speech characteristics modulate such influence based on task type. Chinese character pairs were displayed to native Chinese readers under four auditory conditions: normal Chinese speech, phonotactically legal but meaningless speech, spectrally-rotated speech (i.e. meaningless sound with no accessible phonological form), or silence. Participants were tasked with determining whether the presented character pair shared the same meaning (semantic judgment), or the same initial phoneme (phonological judgment). Participants performed better and faster in the semantic than in the phonological judgment task. Phonological properties of meaningless speech prolonged participants' reaction times in the phonological but not the semantic judgment task, whilst the semantic properties of speech only delayed reaction times in the semantic judgment task. The results indicate that background speech disrupts lexical processing, with the nature of the primary task affecting the extent of phonological and semantic disruption.

A substantial amount of research has demonstrated that background speech that is to-be-ignored, is disruptive to reading (e.g. Bell et al., 2008; Hyönä & Ekholm, 2016; Martin et al., 1988; Meng et al., 2020; Sörqvist et al., 2010). Lexical identification is, arguably, the most basic and fundamental process in reading. The present study was, therefore, designed to examine possible disruption effects by particular properties of background speech on lexical identification of individual words.

Two main alternative theories that seek to explain how background speech disrupts text processing cleave on the distinction between interference-by-content and interference-by-process (Marsh et al., 2008a, 2009). According to the interference-by-content account, disruption arises due to similarity in content between background speech and visually-attended text. It holds that speech stimuli can automatically gain access to

the same representational space as the recoded visual text, thereby interfering with the maintenance and retrieval of visual information being processed (see Salamé & Baddeley, 1982, 1986, for a representative account based on the Working Memory model). Accordingly, this account predicts that the magnitude of disruption is related to the degree of phonological (e.g. Salamé & Baddeley, 1982, 1986) or semantic (e.g. Oberauer & Lange, 2008) similarity in content between background speech and visual text. Some research has lent support to this account. For example, Bell et al. (2008) showed that participants' typed prose recall of propositions from a visually-presented extract of a fairy tale was impaired by the presence of meaningful speech compared to meaningless (reversed) speech. Moreover, semantically related speech—an excerpt from the same fairy tale—as compared with unrelated speech—a portion of an

unrelated fairy tale—produced additional disruption to prose recall performance. However, in contrast, Hyönä and Ekholm (2016) found that speech that was constructed from the text to-be-read did not disrupt reading more than speech constructed from a different, semantically unrelated text, questioning the view that disruption of text processing occurs due to shared semantic content.

The contrasting theory, the interference-by-process account, specifies that auditory distraction occurs due to a conflict between similar processes activated by the focal task and task-irrelevant speech and that this occurs regardless of similarity in content (Jones & Tremblay, 2000; Macken et al., 1999; Marsh et al., 2009). This account emerged to explain the disruptive impact of background sound in the irrelevant sound paradigm whereby 6–8 verbal items (e.g. digits) are to be recalled in strict serial order (the irrelevant sound effect; Colle & Welsh, 1976; Salamé & Baddeley, 1982). The interference-by-process account holds that the irrelevant sound effect results from a clash between the deliberate process of seriating the to-be-remembered items via serial rehearsal and the similar process of seriating (i.e. ordering) sound sequences via the obligatory, preattentive process of streaming (see Bregman, 1990). This accounts for why the irrelevant sound effect does not result from phonological (Jones & Macken, 1995) or semantic similarity (e.g. Buchner et al., 1996) between to-be-remembered and to-be-ignored items. There are some instances in which the post-categorical, lexical-semantic properties of speech have been shown to modulate the degree of disruption sound produces to serial recall. For example, valent words (Buchner et al., 2004; Marsh et al., 2018) and taboo words (Rettie et al., 2023; Röer et al., 2017) are more disruptive to serial recall than neutral words. However, these post-categorical effects emerge for tasks that do not require serial order processing (the missing-item task; Marsh et al., 2018). Thus they appear to reflect stimulus-specific attentional diversion that occurs independently of the processes brought to bear on the focal task (Marsh et al., 2018). On this evidence, the expression of post-categorical effects of auditory distraction within the context of the irrelevant sound paradigm, is underpinned by a mechanism distinct from that which underlies the disruption produced by successive changes within an auditory stream (i.e. "interference-by-process").

The explanatory scope of the interference-by-process account has been extended beyond the irrelevant sound paradigm to tasks that tap semantic processing (Marsh et al., 2008a, 2009). On the interference-by-process account, text processing is suggested to be impaired as a result of a conflict between deliberate processes engaged in the focal task and non-deliberate, automatic processing of the meaning and phonological form of speech sounds (e.g. Meng et al., 2020). Thus far, however, only a small number of investigations into the effects of task demands on auditory distraction have been reported. Marsh and his colleagues assessed auditory distraction for recall of semantic category-exemplars and showed that disruption due to meaningfulness of speech, and the semantic similarity between visual memoranda and irrelevant speech (the between-sequence semantic similarity effect) arose only when instructions emphasised recall by category (Marsh et al., 2009) or free-report (Marsh et al., 2008a) rather than by serial order (for an analogous effect for between-sequence phonological similarity, see Marsh et al., 2008b). Similarly, Marsh et al. (2024) demonstrated a between-sequence semantic similarity effect on correct recall of visual memoranda when participants were oriented to deep (semantic) features of to-be-remembered category-exemplars, but not shallow (orthographic) features.

Vasilev et al. (2019) found comprehension question difficulty modulated disruption of paragraph reading by meaningful speech such that disruption was larger in an easy compared to a difficult question condition. Meng et al. (2020) observed that the meaning of speech was only disruptive when participants were asked to read a sentence and form a judgement as to whether it made sense, but had no influence when participants were required to read sentences to detect a non-character. It should be apparent that, according to the interference-by-process account, the particular properties of speech that cause interference during text processing are not fixed; rather, the characteristics that will lead to interference will depend on the precise nature of the focal task.

To our knowledge, studies investigating auditory distraction effects on isolated lexical identification are lacking. To reiterate, the absence of such studies represents motivation for the current experiments. However, whilst there are few studies assessing distraction in isolated word identification, there has been a considerable amount of research assessing the vulnerability of lexical identification during sentence or passage reading to auditory distraction. Some of these studies have suggested that irrelevant speech does interfere with lexical identification of words during reading, showing that background speech caused longer gaze durations (Cauchard et al., 2012), and longer first-pass progressive fixation times (Hyönä & Ekholm, 2016), as well as delayed lexical frequency effects on first fixation duration (Yan et al., 2018)

compared with silent reading. In contrast, other studies have failed to find such effects or have shown mixed effects. For example, Zhang et al. (2018) investigated how exposure to music (that contained lyrics and, thus, was meaningful) affected passage reading. Zhang et al. observed no significant differences between a background music and a silence condition during reading for gaze duration, first-pass reading time, and word skipping rate, all eye movement measures that are usually taken to reflect lexical and early linguistic processes. However, with a multiple regression analysis, Zhang et al. (2018) found that gaze duration on low- but not high-frequency words was less predictable from word length, suggesting disrupted sublexical processing under music exposure at least for words of low frequency. Further, Vasilev et al. (2019) reported no significant disruption for first fixation duration nor gaze duration, and they also found a normal word frequency effect when individual sentences were read under background speech conditions relative to silence. However, somewhat surprisingly, when passages that contained a greater amount of text content were read, they did observe disruptive effects in first-pass reading measures. Vasilev et al. suggested that this disruption of first-pass paragraph reading may have arisen due to the longer texts content causing readers increased difficulty in maintaining sustained attention through longer periods of reading. Clearly there is some inconsistency amongst auditory distraction studies investigating lexical identification in natural reading, though it is certainly the case that some research has shown that speech may be disruptive to word identification under some circumstances.

As noted earlier, the present study was motivated by a lack of studies examining auditory distraction effects on lexical processing of isolated words. We, therefore, adopted two lexical judgment tasks (following Chiu et al., 2016), one in which participants were instructed to judge whether two Chinese characters shared the same meaning, and the other in which they were required to judge whether the characters shared the same initial phoneme. Given that these tasks examine aspects of lexical processing in the absence of most other linguistic processes that occur during natural reading, it is possible that clearer and less ambiguous auditory distraction effects might be apparent. Furthermore, tasks involving isolated word processing do not require participants to maintain attention to processing over extended passages of text, and therefore, presumably, they are less susceptible to effects driven by attentional failures (cf. Vasilev et al., 2019). Also, since our lexical processing tasks required an explicit judgement in respect of meaning or phonology, we assumed that participants would almost

certainly engage in semantic processing or phonological processing, respectively in order to complete the task. Furthermore, both our lexical judgment tasks required that participants retain the semantic or phonological codes of the two characters in working memory in order that they might be able to form a decision as to their relatedness. Arguably, such memory encoding might likely not occur during natural reading given that no comparative linguistic judgment is required.

In the current experiment, we presented our visual stimuli in four different auditory distraction conditions: normal Chinese speech, phonotactically-legal meaningless speech, spectrally-rotated speech and silence. We adopted variants of spoken Chinese as background sound stimuli because we tested Chinese-speaking participants, and our visual stimuli were Chinese characters. Phonotactically-legal but meaningless speech (PL-MLS), that is, a speech stream comprised of syllables that preserve the phonetic structures of Chinese speech but for which there are no corresponding real characters. PL-MLS has rarely been used for auditory distraction in previous studies, however, we felt the development and use of such a distractor stimulus was important to allow us to determine whether syllabic content in the absence of meaningful words might be sufficient to produce disruptive lexical processing effects. We note that most meaningless speech stimuli adopted in previous studies have taken the form of foreign speech (e.g. Hyönä & Ekholm, 2016; Martin et al., 1988; Vasilev et al., 2019), reversed speech (Jones et al., 1990) or spectrally-rotated speech (Sörqvist et al., 2012), with most such stimuli differing from participants' native speech with respect to phonetic structure. Of course, such meaningless speech stimuli differ from PL-MLS that we adopt here as they contain few or no accessible phonological properties of the natural speech of the participants. As a consequence, it is possible that such stimuli might cause participants to engage in only limited, non-deliberate phonological processing of the speech sound. Empirical evidence for this comes from studies in second-language learning that have consistently shown that second-language learners' phonological awareness scores increase significantly and steadily over time and are influenced by their language proficiency (Gao & Gao, 2005; Mullady-Dellicarpini, 2005; Sakuma & Takaki, 2018). Thus, it appears that participants are not able to engage in phonological processing of speech with an unfamiliar phonological form to the same extent as they can with native speech. In that situation, commonality of phonological content or process in relation to the speech stimuli and visual text will very likely be minimal and this may be a reason why previous studies failed to consistently observe disruption effects of meaningless

speech on reading (e.g. Martin et al., 1988; Vasilev et al., 2019; Yan et al., 2018). It was for these reasons that we used PL-MLS as meaningless speech stimuli, preserving the phonetic structure of native (Chinese) speech, to test whether evidence of phonological distraction on reading may be observed when the stimuli allow for more accessible phonological processing. In addition to PL-MLS, we also adopted a spectrally-rotated speech (SRS) noise control condition. The characteristics of spectrally-rotated speech and original Chinese speech are quite comparable in terms of intonation, rhythm, and the duration of pauses between words and sentences, but spectrally-rotated speech is semantically and phonologically inaccessible to Chinese speakers. The inclusion of these three speech conditions and silence (to assess undisrupted, ceiling performance processing), as a control condition allows for systematic examination of the influence of semantic and phonological properties of speech on semantic and phonological similarity judgment performance.

To sum up, the present study investigated interference effects of background speech on lexical processing associated with isolated words and in the absence of additional linguistic processing that occurs during natural reading. More specifically, we explored whether lexical task demands modulate the magnitude of disruption produced by irrelevant speech. We compared the effects of semantic and phonological properties of speech on two lexical judgment tasks, one requiring semantic and the other requiring phonological processing. By instructing one group of participants to decide whether a pair of visually presented characters shared the same meaning, and another to decide whether the identical set of character pairs shared the same initial phoneme, we assessed the degree to which different dominant focal processes were impacted by our different irrelevant speech manipulations. The interference-by-content account stipulates that content similarity between the speech and the visually presented characters will determine the magnitude of disruption. Thus, the interference-by-content account predicts that disruption by irrelevant speech should occur regardless of task instruction. In contrast, the interference-by-process account predicts an interaction between task instruction and background auditory stimuli, since it supposes that disruption will occur due to the extent that the background sound and visual stimuli draw on similar processes. That is, the semantic properties of irrelevant speech should be more disruptive in a semantic judgment task than in a phonological judgment task; and conversely, the phonological properties of irrelevant speech should be more disruptive in the phonological judgment task than in the semantic judgment task.

Method

Participants

Sixty-four undergraduate students (mean age = 20.5 years, SD = 2.2; 52 females) recruited from Tianjin Normal University were randomly assigned to one of two between-participant groups: semantic judgment vs. phonological judgment tasks (i.e. 32 in each). A between-participants design was adopted to avoid any potential task-transfer contamination effects. All participants reported normal or corrected-to-normal vision, normal hearing and were native Chinese speakers. Participants were rewarded with gifts (such as data cables, liquid soap, 12-color painting sticks, or sketchbooks) for their participation in the experiment. The research received ethical approval from the Research Ethics Committee at Tianjin Normal University (ID: APB20180402). All participants provided electronic informed consent.

Apparatus

A ThinkPad notebook was used to run this experiment. The experimental procedure was programmed and presented in E-prime 2.0 software. The visual Chinese character stimuli were presented on a 14-inch screen with a resolution of 1,920 × 1,080 pixels and a refresh rate of 60 Hz. At 50 cm viewing distance, each character subtended 1.1°. The participant's head was kept immobile by using a head and chin rest.

Materials

Visual stimuli

Each participant responded to 368 character pairs in the formal experiment, including 240 experimental trials, 120 filler trials and eight practice trials (i.e. the first two trials after each change of background sound). The two characters in each experimental trial had no relationship in respect of their semantics or phonology. The experimental trials were identical under the semantic judgment task and the phonological judgment task. But the filler trials differed between the two tasks as the second character in each character pair changed. Specifically, in the semantic judgment task, the character pairs in all filler trials had the same meaning but different initial phonemes; while in the phonological judgment task, the character pairs in all filler trials shared the same initial phoneme but differed in meaning. Examples of character pairs in the experimental and filler trials are presented in Table 1.

Table 2 provides the number of strokes, single-character word frequencies and character frequencies of the second characters in the filler trials based on the

SUBTLEX-CH database (Cai & Brysbaert, 2010), which were matched between two tasks (all *t*s < 1.48, all *p*s > 0.14).

Prior to the formal experiment, there were eight practice trials. Half of the practice trials required a YES response, and half a NO response.

Auditory stimuli

Meaningful speech (MFS) was Chinese narrative taken from China Central Television's evening news broadcast. We used phonotactically-legal meaningless speech as meaningless speech (MLS) stimuli, which was created according to the following steps: First, identify the Pinyin of each Chinese character in the MFS, then retain the initial phoneme and tone of each Pinyin, but replace its rime with an alternative rime to make a spliced Pinyin with regular phonetic structure but no corresponding real character (e.g. Pinyin of Chinese characters 好久不见 is /hao3 jiu3 bu2 jian4/. Based on this, the recombined Pinyin for PL-MLS would be /hing3 jua3 bou2 juang4/).[1] MFS and MLS were recorded in the same adult female voice and sampled with a 16-bit resolution, at a sampling rate of 44.1 kHz using Audacity 2.1.3 software. Spectrally rotated speech (SRS) noise control was created by using Matlab, in which the spectrum of MFS was low-pass filtered at 3.8 kHz and then inverted around 2 kHz (as in Scott et al., 2009). All sounds were diotically delivered via headphones (Newmine MX660), and continuously presented during the entire block in a given irrelevant sound condition. The intensity of three types of speech was 58–72 dB (A). The ambient level for the silent condition was 45 dB(A). All the auditory stimuli were of sufficient duration (no less than 20 min) to extend over the full period that the participants spent judging the character-pairs.

Design

A 2 × 4 mixed design was employed with task (semantic judgment vs. phonological judgment) as a between-participants factor and background sound (MFS vs. MLS vs. SRS vs. silence) as a within-participants factor. The character pairs were divided into four blocks, each consisting of 60 experimental trials, 30 filler trials and two practice trials. The order of the four background sounds was counterbalanced across participants. Thus, each block was presented under each sound condition an equal

number of times across participants. The experimental and filler trials in each block were presented randomly.

Procedure

The start of each trial was signalled by a 300-ms fixation cross presented in the centre of a CRT display. Following this, there was a 300-ms blank interval prior to stimulus presentation. The two characters for a trial were then displayed simultaneously on the same horizontal line of the screen, with one character positioned at the centre of the screen where the fixation cross had appeared and the other character situated to the right of that character. The distance between the centres of the two characters was 2.4°. The characters remained in view either until the participant pressed a response key or until 5 s had elapsed. There was a 2-s interval before the start of the next trial.

Participants were instructed that on each trial they would be presented with two characters and that their task was to decide, as quickly and as accurately as possible, whether the two characters shared the same meaning or the same initial phoneme under the semantic judgment task and the phonological judgment task, respectively. They were asked to ignore the background speech and concentrate only on the task decision. If two characters on a trial shared the same meaning or the same initial phoneme, the participants were instructed to press a YES key; otherwise, they were instructed to press a NO key (please see Table 1). Reaction times (RTs) were recorded from the onset of the characters until the participants responded. The participants were instructed to keep their index fingers resting one on each key to achieve fastest RTs. The experiment lasted approximately 30 min.

Analysis

Data from both the experimental and the filler trials were analysed. Note that although the second character of each filler stimulus pair in each trial differed under the two tasks, their properties were closely matched (see again Table 2), meaning that analyses of data from these trials are very likely valuable and meaningful. Data from the practice trials were discarded.

We undertook analyses of judgment accuracy and reaction times (RTs). RT analyses were performed with linear-mixed effects models and run with the lme4 package (Bates et al., 2015), available in the R

---

[1]Due to differences in experience with alternative dialects, many participants cannot distinguish between the blade-alveolars (i.e. *z, c, s*) and the corresponding retroflexes (i.e. *zh, ch, sh*) when speaking and listening to Chinese. Therefore, whenever a Pinyin with initial phoneme of z/zh, c/ch, or s/sh did not have a corresponding real word (e.g. /*shong3*/), but its corresponding blade-alveolar or retroflex sound (/*song3*/) did correspond to one or more real characters (e.g. 耸[/*song3*/, *towering*], 悚[/*song3*/, *terrified*]), then the combined Pinyin (/*shong3*/) was not selected as a replacement Pinyin.

Table 1. Example trials used in the two tasks.

Task	Trial type	Example	Meaning	Pinyin[2]	Correct response
Semantic judgment	Experimental	沸—歧	Boiling-different	fei4-qi2	NO
	Filler	频—屡	Frequent-frequent	pin2-lü3	YES
Phonological judgment	Experimental	沸—歧	Boiling-different	fei4-qi2	NO
	Filler	频—痞	Frequent-ruffian	pin2-pi3	YES

Table 2. Properties of the second characters of the filler trials under two tasks.

Task	Example	Number of strokes		Single-character word frequency (per million)		Character frequency (per million)	
		M	SE	M	SE	M	SE
Semantic judgment	屡	9.91	0.28	26.40	4.73	111.79	16.05
Phonological judgment	痞	9.75	0.26	23.02	3.66	103.46	15.95

environment (R Core Team, 2018). Generalised linear mixed models (GLMM) were used to analyse accuracy. For each variable, a model was specified with participants and items as crossed random effects, with task and background sound as fixed factors. Four successive difference contrasts were set up to analyse effects across experimental conditions; for effects of semantic meaningfulness (MFS vs. MLS), phonological properties of speech (MLS vs. SRS), acoustic properties of speech (SRS vs. silence) and overall speech (MFS vs. silence). Regression coefficient estimates (b), standard errors (SE), t-values (z-value for the accuracy) and effect sizes (d) are reported. We first ran a full random structure for participants and items. If the initial model failed to converge then the random structure was incrementally trimmed, beginning with the items level. RT data but not accuracy of judgment data were log-transformed prior to analysis. Separate analyses were also performed for each task to tease apart significant interactions.

Results

Experimental trials

Twenty-five trials were dropped because of null response within the 5-s time limit (0.2%). Reaction time data were excluded if (a) a response was not correct (2.2%); (b) a value was more than 3 standard deviations above the mean for each participant and each condition (1.4%); (c) a trial was disturbed due to an irrelevant activity (e.g. sneeze, cough, etc.) during a trial (< 0.01%). The mean error rate and mean correct RT for each condition are shown in Table 3 and Figure 1, respectively.

The analysis of error rates yielded a significant main effect of task (b = 1.12, SE = 0.35, z = 3.19, d = 0.05). Participants made more errors when making phonological compared with semantic judgments (2.5 vs. 1.8%), though the overall error rates were very low. The effect of sound condition and the interaction between task and sound on error rates was not significant (all zs < 1.26). Clearly, the participants were able to perform the tasks well.

The results from the LMMs for RTs are summarised in Table 4. In the analysis of RTs, robust main effects of task and sound were observed. In relation to task, phonological judgments were more difficult to make than semantic judgments (1,804 vs. 1,100 ms). In relation to effects of sound, RTs were longer indicating larger disruption to lexical judgments under MFS compared to MLS conditions, and similarly, disruption was larger under these two conditions than under SRS conditions. Also, the SRS condition and silence condition did not differ significantly. The two-way interactions between task and sound (excluding SRS vs. silence), in which we were most interested, were significant. Two sets of separate analyses were conducted, one for each of the two tasks (see Table 5).

For the semantic judgment task, separate analyses showed no significant interference from phonological properties of speech (MLS vs. SRS). However, there were significant differences between MFS and MLS, and MFS and silence. For the phonological judgment task, interestingly, MFS and MLS increased RTs to an equal degree. Whilst other comparisons, including MLS and SRS, and MFS and silence, showed significant differences.

From these analyses on the experimental trials, we can summarise that the participants who identified characters for meaning had higher accuracy and shorter reaction time in comparison to the participants who

---

[2]Pinyin, is an alphabetic system that employs the alphabet letters to transcribe the exact pronunciation of a Chinese character, including its lexical tone (Lin et al., 2010; Rayner et al., 2012; Wang & Andrews, 2021; Xu et al., 1999; Zhou & Perfetti, 2021). It is important to note that Chinese lacks a productive letter-sound mapping system, and therefore Chinese characters do not explicitly encode their pronunciation. Instead, character pronunciation must be memorized. To aid in this process, primary schools in the Chinese mainland teach the Pinyin system in first grade. As shown in Table 1, /fei4/ is the Pinyin representation of the character 沸, a syllable in which the segments are pronounced /fei/, produced with Tone 4.

Table 3. Mean error rates (%), broken down by task and sound condition for the experimental trials.

Task	MFS	MLS	SRS	Silence
Semantic judgment	1.5 (0.3)	2.3 (0.3)	1.6 (0.3)	1.7 (0.3)
Phonological judgment	2.1 (0.3)	2.5 (0.4)	2.6 (0.4)	3.0 (0.4)

Note: The standard errors are shown in parentheses.

were asked to identify the phonetic structure of characters at the phoneme level. This is consistent with the previous studies examining the relative time course of semantic and phonological activation in reading Chinese, which supports the suggestion that in Chinese reading, semantic information in the lexicon is activated at least as early and just as strongly as phonological information (Chen & Peng, 2001; Chen et al., 2003; Shen & Forster, 1999; Zhou & Marslen-Wilson, 2000, 2009). More importantly, there were reliable interactions between task and background sound: The semantic properties of speech (MFS vs. MLS) increased reaction time when participants were engaged in semantic processing, but not when engaged in phonological processing. In contrast, the phonological properties of speech (MLS vs. SRS) increased reaction time exclusively when participants were engaged in phonological processing. These results indicate that the disruptive effects of background speech on lexical processing are

Table 4. Output from the linear-mixed effects models for reaction time for the experimental trials. Significant effects are marked in bold.

Effect	b	SE	t	d
Intercept (grand mean)	7.17	0.03	**247.2**	/
Task (semantic judgment vs. phonological judgment)	−0.52	0.06	**−9.00**	−1.19
Sound (MFS vs. MLS)	0.07	0.01	**11.14**	0.12
Sound (MLS vs. SRS)	0.07	0.01	**10.80**	0.17
Sound (SRS vs. silence)	−0.04	0.01	−0.68	/
Sound (MFS vs. silence)	0.13	0.01	**21.25**	0.29
Task × Sound (MFS vs. MLS)	0.14	0.01	**10.86**	/
Task × Sound (MLS vs. SRS)	−0.11	0.01	**−8.78**	/
Task × Sound (SRS vs. silence)	0.01	0.01	0.68	/
Task × Sound (MFS vs. silence)	0.03	0.01	**2.76**	/

modulated by both the nature of focal task (i.e. the type of processing in which the participant was engaged) and the linguistic properties of speech sounds. Next, we will consider data on the filler trials to examine the effects of task and background speech on lexical processing.

Filler trials

Eleven filler trials were dropped because of a null response within the 5-s time limit (0.1%). RTs with incorrect responses (8.5%), and RTs that differed by more
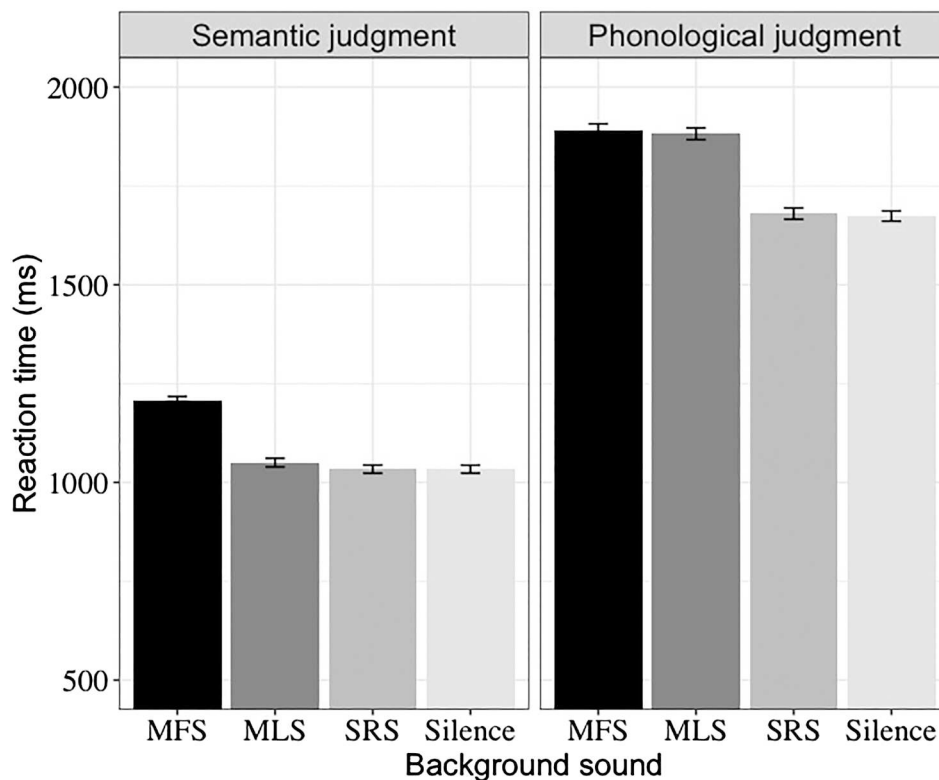

Figure 1. Mean reaction times for the different background sound conditions, broken down by task. Error bars represent the standard error of the means. MFS = meaningful speech; MLS = meaningless speech; SRS = spectrally-rotated speech.

Table 5. Simple effect analysis of the interaction between Task and Sound for the experimental trials. Significant effects are marked in bold.

Effect	Semantic judgment				Phonological judgment			
	b	SE	t	d	b	SE	t	d
Sound (MFS vs. MLS)	0.14	0.01	**15.06**	0.31	0.00	0.01	0.42	/
Sound (MLS vs. SRS)	0.01	0.01	1.40	/	0.12	0.01	**14.89**	0.32
Sound (MFS vs. silence)	0.15	0.01	**16.46**	0.36	0.12	0.01	**14.18**	0.35

Table 6. Mean error rates (%) for filler trials (*SE*s in parentheses).

Task	MFS	MLS	SRS	Silence
Semantic judgment	9.1 (0.9)	10.9 (1.0)	9.9 (1.0)	9.1 (0.9)
Phonological judgment	7.1 (0.8)	7.6 (0.9)	7.4 (0.8)	7.0 (0.8)

Table 7. Output from the linear-mixed effects models for reaction times for filler trials. Significant effects are marked in bold.

Effect	b	SE	t	d
Intercept (grand mean)	7.05	0.03	**271.4**	/
Task (semantic judgment vs. phonological judgment)	−0.40	0.05	**−8.15**	0.94
Sound (MFS vs. MLS)	0.05	0.01	**5.79**	0.09
Sound (MLS vs. SRS)	0.05	0.01	**5.33**	0.14
Sound (SRS vs. silence)	0.01	0.01	0.70	/
Sound (MFS vs. silence)	0.10	0.01	**11.87**	0.25
Task × Sound (MFS vs. MLS)	0.10	0.02	**5.96**	/
Task × Sound (MLS vs. SRS)	−0.12	0.02	**−7.07**	/
Task × Sound (SRS vs. silence)	0.01	0.02	0.40	/
Task × Sound (MFS vs. silence)	−0.01	0.02	−0.73	/

than 3 standard deviations from the mean for each participant and each condition (1.5%) were eliminated from analysis. Mean error rates and mean correct RTs for filler trials are presented in Table 6 and Figure 2, respectively.

Analysis of error rates showed a main effect of task (*b* = −0.35, *SE* = 0.13, *z* = −2.65, *d* = 0.09). Error rates were significantly greater for semantic judgments than for phonological judgments (9.8 vs. 7.3%), a contrast to the error data pattern for experimental trials. This result was likely due to a response bias caused by the

reduced number of filler trials (YES response) relative to experimental trials (NO response). We will return to this issue in the Discussion. No other significant effects were found (all *z*s < 1.41).

The results from the LMMs for RTs of filler trials are summarised in Table 7. As with the experimental trials, mean RTs were significantly faster for semantic judgments than for phonological judgments (1,001 vs. 1,481 ms). Also, MFS increased RTs to a greater degree than did MLS, while SRS did not impair performance markedly compared to silence. The interactions between task and sound (MFS vs. MLS; MLS vs. SRS) were significant. Two sets of separate analyses are
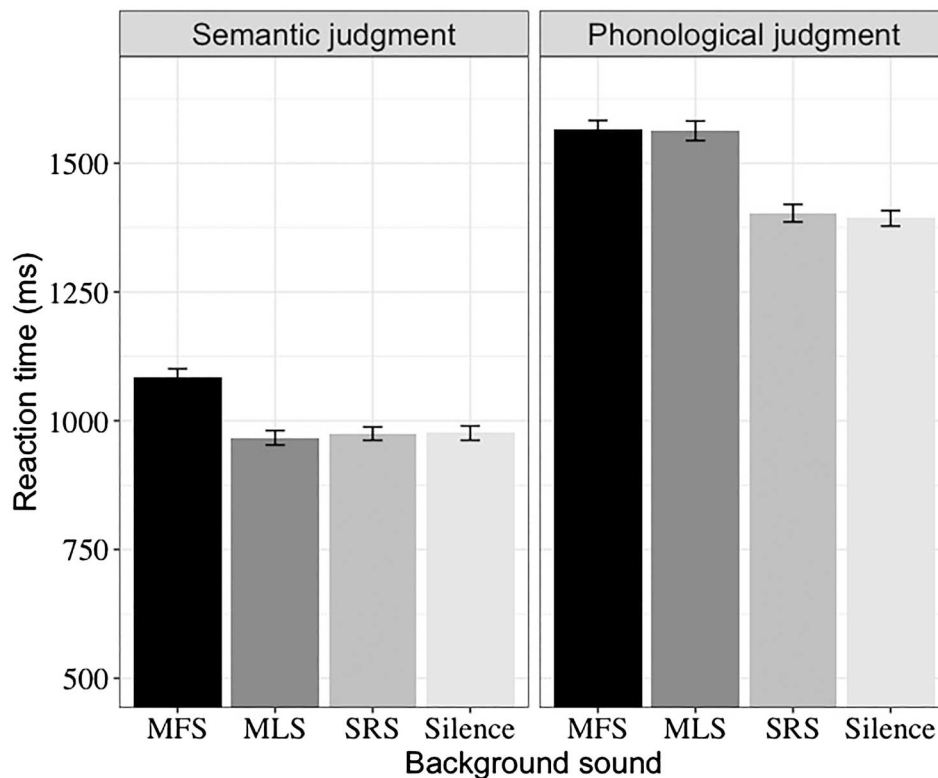

Figure 2. Mean reaction times of filler trials for the different background sound conditions, broken down by task. Error bars show the standard error of the means. MFS = meaningful speech; MLS = meaningless speech; SRS = spectrally-rotated speech.

Table 8. Simple effect analysis of the interaction between Task and Sound for filler trials. Significant effects are marked in bold.

Effect	Semantic judgment				Phonological judgment			
	b	SE	t	d	b	SE	t	d
Sound (MFS vs. MLS)	0.10	0.01	**8.09**	0.26	−0.00	0.01	−0.04	/
Sound (MLS vs. SRS)	−0.01	0.01	−1.08	/	0.11	0.01	**9.64**	0.30

presented in Table 8. MFS increased RTs compared with MLS for the semantic judgment task, but not for the phonological judgment task. Whilst MLS increased RTs compared with SRS for the phonological judgment task, but not for the semantic judgment task.

To summarise the findings from the filler trials, error rates were lower and reaction times were longer when participants were required to judge whether two isolated characters shared the same initial phoneme than when they were required to judge whether the two characters shared the same meaning. More interestingly, the semantic properties of speech (MFS vs. MLS) exclusively delayed semantic judgments, whereas the phonological properties of speech (MLS vs. SRS) were only disruptive to phonological judgments. Overall, these results, alongside the results from the experimental trials, demonstrate that the extent to which distractor speech exerts an influence over lexical processing with isolated characters is determined by the properties of the speech comprising that distractor in relation to the nature of processing required for the focal task.

Discussion

The present study was conducted to examine disruption to lexical processing due to different properties of background speech under different task instructions. Results suggested that the effect of background speech on lexical processing appears to be process—rather than content-driven. In comparison with silence, only meaningful speech (i.e. normal Chinese speech) significantly increased reaction times in a semantic judgment task, whereas both meaningful and meaningless speech produced a comparable increase in participants' reactions times when the task required a phonological judgment. These results provide support for the interpretation of auditory distraction on lexical processing as being process-based.

Previous auditory distraction studies have increased our understanding of the nature of the impact of irrelevant speech on task performance and have shed light on the role played by focal task processes in modifying the magnitude of any disruption effect. However, those studies mainly focused on short-term memory or complex reading tasks, and few of them have examined distraction effects for processing of isolated words. For example, Marsh et al. (2009) reported distraction by

irrelevant speech on recall of category-exemplars and revealed that the disruptive effects of meaningful speech arose when participants adopted a retrieval strategy based on semantic-categorization but not when it was based on seriation. Furthermore, Marsh et al. (2009) found that meaningful speech reduced the adoption of a semantic-organization strategy in a free recall task, as indexed by a diminution in the propensity to cluster recalled items by category. These results suggested that meaningful speech caused disruption to the strategy or process underpinning the focal task. Similarly, Meng et al. (2020) found that disruption in sentence processing by meaningful background speech only occurred when the task required semantic comprehension of the text. When participants were required to scan sentences to identify an orthographically illegal non-character, no such disruption occurred. Follow-up analyses demonstrated strong lexical frequency effects, as indexed by fixation durations, for both tasks thereby ruling out the notion that the non-character detection was immune to disruption by meaningful background speech simply because it did not engage linguistic processing.

The present results align well with these studies in showing significant interactive effects between task instruction and background sound. More specifically, the present results show directionality of effects in relation to task. That is to say, the semantic properties of speech increased participants' reaction times in the semantic judgment task but not the phonological task, whilst the phonological properties of speech (regardless of its meaningfulness) increased participants' reaction times exclusively in the phonological judgment task. To reiterate, these results fit neatly with the interference-by-process account, which stipulates that the degree of auditory distraction that will occur on a particular task is determined jointly by the properties of the auditory stimulus (in this case speech) as well as the nature of the focal task.

The interference-by-content account, to us, provides a less compelling explanation of the results reported here. In this study, while task instructions differed, background sounds and visual materials (experimental trials) remained consistent across tasks. This meant that there was no difference in content similarity between auditory and visual materials across tasks. Therefore, given that background sounds had different effects for reaction

times on the two Chinese character recognition tasks, this suggests that these effects are due to differences in task processing. In sum, our interpretation of these effects is that shared content between background speech and text does not determine the magnitude of disruption caused by irrelevant speech, but instead, the nature of the primary task and the visual and cognitive processing associated with that task plays a significant role. However, it must be noted that the degree of semantic or phonological content similarity between speech and visual text was not directly manipulated in the present experiment. That is to say, whilst the current results do provide evidence for effects of primary task and process, the experiment did not afford the opportunity to directly observe differential content effects. Clearly, to deliver a more robust assessment of the interference-by-content account, such experimental conditions would be necessary.

In fact, several studies have already questioned the interference-by-content account by directly manipulating the content similarity of visual and auditory materials using other tasks. For example, studies with short-term memory tasks have shown that semantic or phonological similarity between irrelevant speech items and to-be-remembered visual items has little, if any, impact when participants are required to recall items in serial order (Buchner et al., 1996; Jones & Macken, 1995; LeCompte & Shaibe, 1997). However, content similarity has significant impact if free recall of semantic or rhyme category-exemplars is required (Marsh et al., 2008b, 2009). Also, Neely and LeCompte (1999) found a disruptive effect of semantic similarity in content between visual words and words presented in background speech during serial recall, but this effect was much smaller than that observed in free recall of category-exemplars. More recently, Marsh et al. (2024) reported that the free recall of visually-presented target items was more disrupted by to-be-ignored auditory items from the same semantic category than from a different semantic category. Note, though, that this between-sequence semantic effect only occurred in a task that required words to be processed to a relatively deep level (a pleasantness-rating task), but not in a task that required relatively less depth of processing (a vowel-counting task). Taken together, these studies suggest that the presence and magnitude of between-sequence content similarity effects is influenced by the nature of processing associated with primary task processing. Again, to us, these results complement the current findings and suggest that such effects may not be well explained within the interference-by-content account. Regardless, what is clear from the current findings is that the nature of visual and cognitive

processing associated with the primary task plays an important role in the auditory distraction effect.

The present study revealed significant main effects of task on both error rates and RTs. That is, participants made more errors and took longer when making phonological judgments than was the case when making semantic judgments for the experimental trials. This aspect of the results is consistent with previous findings demonstrating that effects associated with phonological activation deriving from orthographic stimuli are less immediate than effects associated with semantic activation deriving from orthographic stimuli (e.g. Chen & Peng, 2001; Chen et al., 2003; Shen & Forster, 1999; Wang et al., 2021; Zhou & Marslen-Wilson, 2009; but see Tan & Perfetti, 1998). For example, Zhou and Marslen-Wilson (2000) observed strong semantic priming effects in a Chinese character decision task (legal or illegal character) at both short and long SOAs, whilst phonological priming effects were reduced relative to the semantic effects and were observed only at the long SOA. These results suggest that the time to access phonological information associated with an orthographic form is at least as long, and under some circumstances longer, than the time to access semantic information during Chinese character recognition. These results also imply that the recovery of semantic information for Chinese characters does not depend on prior activation of phonological information. Differences in the nature and time course of semantic and phonological activation in Chinese character identification probably arise due to differences in the nature of the relations between orthographic forms and corresponding phonological and semantic representations in this logographic orthography. Relations between orthographic forms and phonological forms are much more arbitrary in logographic languages like Chinese than is the case for more regular languages (e.g. alphabetic languages). Thus, it has been suggested (e.g. Shen & Forster, 1999; Wang et al., 2021; Zhou & Marslen-Wilson, 2000, 2009) that for Chinese, a direct route from orthography to meaning is dominant whereas a phonologically mediated route plays a subsidiary role, and that this might represent a more efficient manner of processing for Chinese character identification. Under this assumption, in the present study it is likely that participants activated character meanings directly from orthography in the semantic judgment task. In contrast, in the phonological judgment task, phonological forms may have either been accessed via the semantic route which would require an additional processing step, or alternatively, via a phonologically mediated route involving irregularity and inconsistency. Moreover, recall that the phonological decision task required

participants to judge whether the two characters share the same initial phoneme, thus an extra step of identifying the initial phonemes of two characters after obtaining their phonological forms was necessary. If this suggestion is correct, it might explain why participants took longer and were more error-prone in the phonological judgment task compared to the semantic judgment task.

Data from the filler trials were almost entirely consistent with the results from the experimental trials. The only notable difference occurred in relation to the main effect of task on error rates, that is, participants made more errors when making semantic judgments than when making phonological judgments for filler trials, the opposite pattern to that obtained for the experimental trials. This effect probably arose due to the difference in the number of experimental trials compared to filler trials. Recall, the ratio of experimental trials to filler trials was 2:1 (60 and 30, respectively under each sound condition). And, an appropriate correct response for an experimental trial was NO, and an appropriate correct response for a filler trial was YES. Consequently, the imbalance in experimental to filler trial ratio, along with inconsistent response patterns, led participants to develop a response bias towards pressing the NO key. As a result, error rates would decrease for experimental trials and increase for filler trials. Evidence to support this suggestion comes from the fact that mean error rates were significantly lower for experimental trials than that for filler trials (2.2 vs. 8.5%). After obtaining this result, we also checked the lexical characteristics of our experimental and filler stimuli. According to the SUBTLEX-CH database (Cai & Brysbaert, 2010), the characters used in our experimental trials had more strokes (10.79 vs. 9.71; $p < 0.001$), lower single-character word frequency (15.86 vs. 294.38; $p = 0.02$), and lower character frequency (45.75 vs. 341.74; $p = 0.003$) than the characters used in the filler trials. If anything, these characteristics should have worked against the pattern of effects for the error rates that we actually obtained, suggesting that the effects were very unlikely due to the lexical characteristics of the experimental and filler stimuli. Consequently, it seems likely that the response bias explanation is the more likely reason for the difference.

Indeed, our task here, in which participants were instructed to judge whether the character pairs share the same meaning, or the same initial phoneme, may be considered in signal detection terms. That is, the filler and experimental trials were like signal and noise, respectively. Participants' judgment criteria can vary depending on the probability of signal occurring in respect of noise and this can produce a response bias (Nevin, 1969; Wixted, 2020). Note also that the magnitude of any response bias that might occur with respect to judgements might differ between the semantic and phonological judgment tasks. For example, it might be argued that semantic judgments are more subjective and thus more susceptible to response bias than phonological judgments which are more objective (and therefore less affected by response bias). To be clear, judging whether, or not, two phonemes agree is a judgement that can be made with more certainty than judging whether two terms have the same meaning because the initial phoneme of a Chinese character is unequivocal and singular, whereas meanings between characters may differ in subtle and nuanced ways. Further, it has been established that developing Chinese readers, such as first-grade students aged 6–8 years old, can proficiently recognise the initial phonemes of Chinese characters very likely due to them having learnt Pinyin (e.g. Lin et al., 2010; Newman et al., 2011).

In contrast, semantic judgments are more subjective, as they require understanding and interpretation of the meaning of the characters, and this is related to language, cultural knowledge, and personal experience. Therefore, even for native speakers, there may be nuanced differences in understandings of the meaning of certain Chinese characters (e.g. Passonneau et al., 2012; Ramsey, 2022). To test our assumption, we undertook analyses of response bias in the semantic and phonological judgment tasks with a nonparametric measure B''[3] (Stanislaw & Todorov, 1999). B'' can range from -1 (extreme bias in favour of *yes* response) to 1 (extreme bias in favour of *no* response). In the semantic decision task, B'' had a value of 0.67, whereas in the phonological decision task, B'' has a lower value of 0.46. These results align with our suggestion of a response bias in favour of *no* responses, and that the bias was greater for the semantic than for the phonological judgments. As noted earlier, such a response bias would serve to decrease error rates in experimental trials, and increase error rates in filler trials, and this effect would be greater for semantic judgments (experimental, 1.8%; filler, 9.8%) than for phonological judgments (2.5 and 7.3%, respectively). In short, there does appear to be some evidence that response bias resulting from an imbalance of experimental and filler trials, differentially influenced semantic and phonological judgments, and

---

[3]The formula for B'' is $B'' = \dfrac{H(1 - H) - F(1 - F)}{H(1 - H) + F(1 - F)}$ when $H \geq F$. H indicates the hit rate, that is, the accuracy rate for signal (i.e. filler) trials. F indicates the false-alarm rate, that is the error rate for noise (i.e. experimental) trials (Stanislaw & Todorov, 1999).

this may provide some explanation for the opposite pattern of effects that we observed in the error data. Of course, further direct research is required to verify this suggestion. The much more important aspect of the results from the filler trials was the significant inter-action between task and sound, with semantic proper-ties of speech solely increasing RTs for the semantic judgment task and phonological properties of speech solely increasing RTs for the phonological judgment task. These results are entirely consistent with an inter-ference-by-process view of auditory distraction whereby disruption is a function of a conflict between similar processes.

Beyond the theoretical implications of our results, the present study also indicated that the adoption of PL-MLS as a meaningless speech distractor stimulus is useful and reasonable for studying the influence of phonological properties of speech in auditory distraction. Compared with alphabetic scripts, the method of constructing background sound material that conforms to the pho-netic structure rules of the native language, but lacks the semantic components of Chinese, is more compli-cated. Specifically, in alphabetic languages like English, a pronounceable but meaningless word list might be simply created by changing a single letter of a word that appears in normal speech (e.g. we can create the nonword LANT by replacing the "D" with a "T" in the real word LAND, see Marsh et al., 2008a). This is because letters in alphabetic scripts like English are the smallest orthographic unit and some letters in words may correspond to a phoneme. Consequently, in English, it is possible to construct nonwords that are still pronounceable (e.g. the nonword LANT has a readily accessible phonological form and is, therefore, very readily pronounceable). However, in Chinese, even though the smallest orthographic unit is a stroke, it is not possible to create meaningless speech by produ-cing non-characters in which one stroke is changed, or the position of the radicals is altered. It is important to understand that all non-characters in Chinese are unpro-nounceable because each stroke that makes up a char-acter has no corresponding phonetic form. Thus, the phonological code of a Chinese character cannot be decomposed based on its constituent strokes. And con-sequently, for native Chinese speakers, meaningless speech with accessible phonological properties must be created based on the Pinyin system. The specific method of creating phonotactically-legal meaningless Chinese speech developed in the present study may, therefore, be valuable to future researchers investigating auditory distraction effects of Chinese speech.

In summary, the experiment reported here is one of very few studies that have examined the effects of irrelevant sound on lexical processing of isolated words (Chinese characters). The results clearly indicate that lexical judgment tasks, like semantic or phonologi-cal judgments are sensitive to disruption from irrelevant sound just as are laboratory-based tasks (e.g. serial short-term memory tasks) or complex natural cognitive pro-cessing tasks (e.g. sentence reading and writing). The pattern of results obtained in the present study is best explained by the interference-by-process account that stresses the importance of similarity in shared proces-sing associated with the focal task and background speech. It appears that processing of information con-veyed by speech is activated quite automatically and this then disrupts processing that is similar in nature and is required for the focal task.

Data availability statement

The data files, R scripts, and stimuli used in the present research are openly available at: https://osf.io/j46mr/

ORCID

Zhu Meng 🄳 http://orcid.org/0000-0002-9530-5055
John E. Marsh 🄳 http://orcid.org/0000-0002-9494-1287
Simon P. Liversedge 🄳 http://orcid.org/0000-0002-8579-8546

References

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

Bell, R., Buchner, A., & Mund, I. (2008). Age-related differences in irrelevant-speech effects. *Psychology and Aging*, *23*(2), 377–391. https://doi.org/10.1037/0882-7974.23.2.377

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. MIT Press.

Buchner, A., Irmen, L., & Erdfelder, E. (1996). On the irrelevance of semantic information for the "irrelevant speech" effect. *The Quarterly Journal of Experimental Psychology Section A*, *49*, 765–779. https://doi.org/10.1080/713755633

Buchner, A., Rothermund, K., Wentura, D., & Mehl, B. (2004). Valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition*, *32*(5), 722–731. https://doi.org/10.3758/BF03195862

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

Chen, B. G., & Peng, D. L. (2001). The time course of graphic, phonological and semantic information processing in Chinese character recognition (I). *Acta Psychologica Sinica*, *33*, 1–6.

Chen, B. G., Wang, L. X., & Peng, D. L. (2003). The time course of graphic, phonological and semantic information processing in Chinese character recognition (II). *Acta Psychologica Sinica*, *35*, 576–581.

Chiu, Y. S., Kuo, W. J., Lee, C. Y., & Tzeng, O. J. L. (2016). The explicit and implicit phonological processing of Chinese characters and words in Taiwanese deaf signers. *Language and Linguistics*, *17*, 63–87. https://doi.org/10.1177/1606822X15614518

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

Gao, L. Q., & Gao, X. L. (2005). A study of foreign students' Chinese phonological awareness development. *Journal of Yunnan Normal University*, *3*, 7–13. https://doi.org/10.16802/j.cnki.ynsddw.2005.03.002

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

Jones, D. M., Miles, C., & Page, C. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory? *Applied Cognitive Psychology*, *4*(2), 89–108. https://doi.org/10.1002/acp.2350040203

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review*, *7*(3), 550–558. https://doi.org/10.3758/BF03214370

LeCompte, D. C., & Shaibe, D. M. (1997). On the irrelevance of phonological similarity to the irrelevant speech effect. *The Quarterly Journal of Experimental Psychology Section A*, *50*(1), 100–118. http://doi.org/10.1080/713755679

Lin, D., McBride-Chang, C., Shu, H., Zhang, Y., Li, H., Zhang, J., Aram, D., & Levin, I. (2010). Small wins big: Analytic pinyin skills promote Chinese word reading. *Psychological Science*, *21*(8), 1117–1122. https://doi.org/10.1177/0956797610375447

Macken, W., Tremblay, S., Alford, D., & Jones, D. (1999). Attentional selectivity in short-term memory: Similarity of process, not similarity of content, determines disruption. *International Journal of Psychology*, *34*(5–6), 322–327. https://doi.org/10.1080/002075999399639

Marsh, J. E., Hanczakowski, M., Beaman, C. P., Meng, Z., & Jones, D. M. (2024). *Thinking about meaning: Level-of-processing modulates semantic auditory distraction,* Submitted for publication.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008a). 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

Marsh, J. E., Vachon, F., & Jones, D. M. (2008b). When does between-sequence phonological similarity promote irrelevant sound disruption? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 243–248. https://doi.org/10.1037/0278-7393.34.1.243

Marsh, J. E., Yang, J., Qualter, P., Richardson, C., Perham, N., Vachon, F., & Hughes, R. W. (2018). Postcategorical auditory distraction in short-term memory: Insights from increased task load and task type. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(6), 882–897. https://doi.org/10.1037/xlm0000492

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*(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

Mullady-Dellicarpini, M. E. (2005). Phonological awareness and adult second language literacy development [Doctoral dissertation]. State University of New York at Stony Brook.

Neely, C. B., & LeCompte, D. C. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory & Cognition*, *27*(1), 37–44. http://doi.org/10.3758/BF03201211

Nevin, J. A. (1969). Signal detection theory and operant behavior: A review of David M. Green and John A. Swets' signal detection theory and psychophysics. *Journal of the Experimental Analysis of Behavior*, *12*(3), 475–480. https://doi.org/10.1901/jeab.1969.12-475

Newman, E. H., Tardif, T., Huang, J., & Shu, H. (2011). Phonemes matter: The role of phoneme-level awareness in emergent Chinese readers. *Journal of Experimental Child Psychology*, *108*(2), 242–259. https://doi.org/10.1016/j.jecp.2010.09.001

Oberauer, K., & Lange, E. B. (2008). Interference in verbal working memory: Distinguishing similarity-based confusion, feature overwriting, and feature migration. *Journal of Memory and Language*, *58*(3), 730–745. https://doi.org/10.1016/j.jml.2007.09.006

Passonneau, R. J., Bhardwaj, V., Salleb-Aouissi, A., & Ide, N. (2012). Multiplicity and word sense: Evaluating and learning from multiply labeled word sense annotations. *Language*

Resources and Evaluation, *46*(2), 219–252. https://doi.org/10.1007/s10579-012-9188-x

Ramsey, R. (2022). Individual differences in word senses. *Cognitive Linguistics*, *33*(1), 65–93. https://doi.org/10.1515/cog-2021-0020

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012). *Psychology of reading* (2nd ed.). Psychology Press.

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rettie, L., Potter, R. F., Brewer, G., Degno, F., Vachon, F., Hughes, R. W., & Marsh, J. E. (2023). Warning—Taboo words ahead! Avoiding attentional capture by spoken taboo distractors. *Journal of Cognitive Psychology. Advance online publication*. https://doi.org/10.1080/20445911.2023.2285860

Röer, J. P., Körner, U., Buchner, A., & Bell, R. (2017). Attentional capture by taboo words: A functional view of auditory distraction. *Emotion*, *17*(4), 740–750. https://doi.org/10.1037/emo0000274

Sakuma, Y., & Takaki, S. (2018). Phonological awareness in EFL elementary school students participating in foreign-(English-) language activities. *Journal of Center for Regional Affairs*, *29*, 5–14.

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 Behavior*, *21*(2), 150–164. https://doi.org/10.1016/S0022-5371(82)90521-7

Salamé, P., & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society*, *24*(4), 263–265. https://doi.org/10.3758/BF03330135

Scott, S. K., Rosen, S., Beaman, C. P., Davis, J. P., & Wise, R. J. S. (2009). The neural processing of masked speech: Evidence for different mechanisms in the left and right temporal lobes. *The Journal of the Acoustical Society of America*, *125*(3), 1737–1743. https://doi.org/10.1121/1.3050255

Shen, D., & Forster, K. I. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language and Cognitive Processes*, *14*(5–6), 429–459. https://doi.org/10.1080/016909699386149

Sörqvist, P., Halin, N., & Hygge, S. (2010). Individual differences in susceptibility to the effects of speech on reading comprehension. *Applied Cognitive Psychology*, *24*(1), 67–76. https://doi.org/10.1002/acp.1543

Sörqvist, P., Nöstl, A., & Halin, N. (2012). Disruption of writing processes by the semanticity of background speech. *Scandinavian Journal of Psychology*, *53*(2), 97–102. https://doi.org/10.1111/j.1467-9450.2011.00936.x

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Tan, L. H., & Perfetti, C. A. (1998). Phonological codes as early sources of constraint in Chinese word identification: A review of current discoveries and theoretical accounts. *Reading and Writing*, *10*(3/5), 165–200. https://doi.org/10.1023/A:1008086231343

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), 1484–1512. https://doi.org/10.1037/xhp0000680

Wang, Q., & Andrews, J. F. (2021). Chinese Pinyin: Overview, history and use in language learning for young deaf and hard of hearing students in China. *American Annals of the Deaf*, *166*(4), 446–461. https://doi.org/10.1353/aad.2021.0038

Wang, Y., Jiang, M., Huang, Y., & Qiu, P. (2021). An ERP study on the role of phonological processing in reading two-character compound Chinese words of high and low frequency. *Frontiers in Psychology*, *12*, 637238. https://doi.org/10.3389/fpsyg.2021.637238

Wixted, J. T. (2020). The forgotten history of signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 201–233. https://doi.org/10.1037/xlm0000732

Xu, Y., Pollatsek, A., & Potter, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 838–857. https://doi.org/10.1037/0278-7393.25.4.838

Yan, G. L., Meng, Z., Liu, N. N., He, L. Y., & 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

Zhang, H., Miller, K., Cleveland, R., & Cortina, K. (2018). How listening to music affects reading: Evidence from eye tracking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(11), 1778–1791. https://doi.org/10.1037/xlm0000544

Zhou, L., & Perfetti, C. (2021). Consistency and regularity effects in character identification: A greater role for global than local mapping congruence. *Brain and Language*, *221*, 104997. https://doi.org/10.1016/j.bandl.2021.104997

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*(5), 1245–1265. https://doi.org/10.1037/0278-7393.26.5.1245

Zhou, X., & Marslen-Wilson, W. (2009). Pseudohomophone effects in processing Chinese compound words. *Language and Cognitive Processes*, *24*(7–8), 1009–1038. https://doi.org/10.1080/01690960802174514