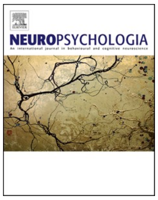
Neuropsychologia 146 (2020) 107562

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: http://www.elsevier.com/locate/neuropsychologia



# The temporal dynamics of first and second language processing: ERPs to spoken words in Mandarin-English bilinguals

1990

Jin Xue [a], Banban Li [a], Rong Yan [b], Jeffrey R. Gruen [c,d], Tianli Feng [e], Marc F. Joanisse [f,g], Jeffrey G. Malins [c,g,h,*]

[a] *University of Science and Technology Beijing, School of Foreign Studies, 30 Xueyuan Road, Haidian District, Beijing, 100083, China*
[b] *Institute of Leadership and Education Advanced Development, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China*
[c] *Yale University School of Medicine, Department of Pediatrics, Yale Child Health Research Center, 464 Congress Avenue, New Haven, CT, 06520, USA*
[d] *Yale University School of Medicine, Department of Genetics, 333 Cedar Street, New Haven, CT, 06520, USA*
[e] *Beijing International Studies University, School of English Language, Literature and Culture, 1 Dingfuzhuan Nanli, Chaoyang District, Beijing, 100024, China*
[f] *The University of Western Ontario, Department of Psychology & Brain and Mind Institute, Western Interdisciplinary Research Building, London, N6A 3K7, Canada*
[g] *Haskins Laboratories, 300 George St. Suite 900, New Haven, CT, 06511, USA*
[h] *Georgia State University, Department of Psychology, P.O. Box 5010, Atlanta, GA, 30302, USA*

## ARTICLE INFO

*Keywords:*
Bilingualism
Spoken word recognition
Phonological competition
Second language processing
Event related potentials

## ABSTRACT

The dynamics of bilingual spoken word recognition remain poorly characterized, especially for individuals who speak two languages that are highly dissimilar in their phonological and morphological structure. The present study compared first language (L1) and second language (L2) spoken word processing within a group of adult Mandarin-English bilinguals ($N = 34$; ages 18–25). Event-related potentials (ERPs) were recorded while participants completed the same cross-modal matching task separately in their L1 Mandarin and L2 English. This task consisted of deciding whether spoken words matched pictures of items. Pictures and spoken words either matched (e.g., Mandarin: TANG2-*tang2*; English: BELL-*bell*), or differed in word-initial phonemes (e.g., Mandarin: TANG2-*lang2*; English: BELL-*shell*), word-final phonemes (e.g., Mandarin: TANG2-*tao2*; English: BELL-*bed*), or whole words (e.g., Mandarin: TANG2-*xia1*: English: BELL-*ham*). Each mismatch type was associated with a pattern of modulation of the Phonological Mapping Negativity, the N400, and the Late N400 that was distinct from those of the other mismatch types yet similar between the two languages. This was interpreted as evidence of incremental processing with similar temporal dynamics in both languages. These findings support models of spoken word recognition in bilingual individuals that adopt an interactive-activation framework for both L1 and L2 processing.

## 1. Introduction

Although first language (L1) and second language (L2) processing are generally thought to share many common neural resources (Higby et al., 2013; Costa and Sebastián-Gallés, 2014; Golestani, 2016), there remains considerable debate concerning the factors that govern the extent of convergence between L1 and L2 processing systems (Clahsen and Felser, 2006; Perani and Abutalebi, 2005). An emerging view suggests that the influence of multiple factors, such as age of acquisition, proficiency, and degree of cross-linguistic similarity between languages, may differ between levels of language processing (Del Maschio and Abutalebi, 2019; Marian et al., 2003), such as phonology or syntax, and even with the same level of processing depending on the routines employed (Sabourin and Stowe, 2008). For instance, phonological and syntactic processing are more susceptible to age of acquisition, whereas lexico-semantic processing is more sensitive to language proficiency (Del Maschio and Abutalebi, 2019). In addition, within the level of syntactic processing, cross-linguistic differences impact processing of grammatical gender to a greater extent than they impact verbal domain processing (Sabourin and Stowe, 2008).

In the domain of spoken word recognition, an outstanding question concerns the extent to which the neurocognitive processes that allow listeners to resolve competition between similarly sounding words are convergent between L1 and L2 when they are highly distinct in their phonological and morphological structure. Motivated by work suggesting that the neurocognitive processes supporting Mandarin spoken

---

* Corresponding author. Georgia State University, Department of Psychology, P.O. Box 5010, Atlanta, GA, 30302, USA.
*E-mail address:* jmalins@gsu.edu (J.G. Malins).

https://doi.org/10.1016/j.neuropsychologia.2020.107562
Received 24 June 2019; Received in revised form 12 July 2020; Accepted 13 July 2020

word recognition might be different than those used to support English spoken word recognition (Zhao et al., 2011), the present study aimed to compare the timing and nature of first and second language spoken word processing within the same group of Mandarin-English bilinguals. By using event related potentials (ERPs) to characterize the neural patterns supporting spoken word recognition within each language, our goal was to offer insights into theories of language processing, especially in the context of bilinguals who learn two languages that are typologically distinct.

### 1.1. *Using ERPs to study processing dynamics*

ERPs provide information about the temporal dynamics of language processing, and offer insights into the neurocognitive processes that precede overt behavioral responses. By time-locking ERP measurements to the point in time at which individuals hear spoken words, different ERP components have been linked to different stages of lexical processing. One well established ERP correlate of lexical processing during spoken word recognition tasks is the Phonological Mapping Negativity (PMN). The PMN, which peaks 230–310 ms in fronto-central sites after the onset of target words (Desroches et al., 2009), is a neural marker for sensitivity to sublexical phonological information (Archibald and Joanisse, 2011), indexing the mapping of acoustic input onto phonemic expectations for both words and nonwords (Newman et al., 2012). The PMN is greater in amplitude when the initial phoneme of a target word mismatches the high cloze probability word in spoken word recognition compared to when the initial phoneme matches expectations (Connolly and Phillips, 1994). For instance, in the sentence "The pig wallowed in the pen", the speech input "pen" is semantically appropriate in context, but induces a PMN due to a mismatch in the initial phoneme from the expected word "mud".

The PMN is independent of the N400, which is largest over centro-parietal sites and peaks around 400 ms after stimulus onset. Unlike the PMN that reflects sublexical processing, the N400 (Kutas and Hillyard, 1980) is attributed to processes related to the integration of word- and/or semantic-level information. The N400 is modulated in semantically anomalous or unexpected contexts (Kutas and Federmeier, 2011). It can also be observed in contexts in which a PMN is not present – for instance, only a delayed N400 is observed for words that are semantically anomalous yet share initial phonemes with the highest cloze probability word (e.g., "The gambler had a streak of bad luggage"; Connolly and Phillips, 1994).

Previous work has also shown that the PMN and N400 can be differentially modulated during word recognition based on the manner in which a target word deviates from a listener's expectation based on the context (Desroches et al., 2009, 2013; Kornilov et al., 2015; Malins et al., 2013, 2014; Malins and Joanisse, 2012). In a typical study, participants are presented with a picture and then hear a matching or mismatching spoken word. The picture sets up an expectation that is subsequently met or violated by the target word (Desroches et al., 2009). By varying the relationship between presented and expected words, this cross-modal picture-spoken word matching paradigm can reveal the time course over which the listener uses incoming auditory information to recognize a spoken word. In Desroches et al. (2009), rhyme mismatches between spoken words and picture items (see "CONE", hear "*bone*") resulted in an increase of the PMN and a reduction of the N400, whereas cohort mismatches (CONE-*comb*) did not modulate the PMN but instead incurred a late increased negativity of the N400 (from 410 to 600 ms, which the authors called the Late N400). Rhyme effects were interpreted to be the consequence of both earlier phonological mismatches (as indexed by the increased PMN) as well as top-down priming of rhyme neighbors that resulted in facilitated recognition of word-final phonemes (as indexed by a reduced Late N400). In contrast, cohort effects were interpreted to be the result of the increased effort required to overcome the effect of misleading bottom-up information, as word-initial phonemes overlapped with phonological expectations and mismatches were not signaled until later on during the unfolding of spoken words (Desroches et al., 2009).

As illustrated by these patterns of responses to cohort and rhyme mismatches, PMN and N400 effects reflect both bottom-up phonological activation as well as top-down selection among lexical representations. This evidence supports interactive activation models such as the TRACE model (McClelland and Elman, 1986), which propose a continuous mapping of speech input onto lexical representations as spoken words unfold, with the extent of activation influenced by the degree of phonological similarity between presented and expected words.

### 1.2. *Phonological analysis in Mandarin versus English*

Although models such as TRACE are well-established based on findings from Indo-European languages, Chinese languages such as Mandarin have phonological features that might result in spoken words being processed using different mechanisms from those captured by continuous mapping models. In English, the phoneme is the basic segmental component for a written word, whereas in morphosyllabic Mandarin, one morpheme (usually one character) corresponds to one syllable. From the standpoint of phonology, Mandarin words generally consist of fewer syllables, and fewer possible patterns of consonant clusters are permitted. For example, syllables are limited to the following structures: consonant-vowel(s)-nasal (CVN; e.g., /kan/ for the character "看"), consonant-vowel(s) (CV; e.g., /ka/ for the character "卡"), vowel(s)-nasal (VN; e.g., /an/ for the character "按"), or vowel(s) (V; e.g., /a/ for the character "阿"). Accordingly, some researchers have argued that more emphasis is placed on the syllable than on the segment across different domains of Chinese language processing, including spoken word recognition (Zhao et al., 2011), visual word recognition (McBride-Chang et al., 2008), and speech production (Chen, O'Séaghdha and Chen, 2016; Wong, Wang, Wong and Chen, 2018).

For instance, in a study by Zhao et al. (2011), participants were asked to judge whether a target picture and a subsequently presented picture belonged to the same semantic category. Between the presentation of the two pictures, participants heard a spoken word that phonologically matched or mismatched the name of the target picture, but they were not required to respond to the spoken word in any way (as opposed to the Desroches et al. (2009) cross-modal picture-spoken word matching task, in which participants were required to actively respond to spoken word stimuli). Zhao et al. (2011) observed that onset (rhyme; e.g., *bi2-li2*) and rime (cohort; e.g., *bi2-bo2*) mismatches modulated N400 amplitudes in a similar fashion despite being signaled at different points in time during the unfolding of Mandarin spoken words. Furthermore, whole-syllable mismatches (e.g., *bi2-ge1*) elicited earlier and stronger N400 effects compared to onset and rime mismatches, which only differed from targets in part of the syllable. Based on this evidence, Zhao et al. (2011) argued that recognition of monosyllabic Mandarin words might rely more on global similarity of the whole syllable structure, or syllable-based holistic processing, as opposed to the phonemic or segment-based incremental processing that is captured by continuous mapping models. That is, partial syllable mismatches are treated equivalently regardless of whether they deviate in onset or rime (or tone) – and thus regardless of whether they are potentially signaled at different points in time – and furthermore whole-syllable mismatches give rise to mismatch effects that are greater in magnitude than mismatches in individual components.

We have previously used the Desroches et al. (2009) ERP cross-modal picture-spoken word matching task to evaluate sensitivity to word-initial (cohort) and word-final (rhyme) processing of Mandarin spoken words in adult native Mandarin speakers (Malins and Joanisse, 2012; Malins et al., 2014) as well as typically developing Mandarin-speaking school-aged children (Malins et al., 2014). In contrast to the Zhao et al. (2011) results, across all studies we observed that patterns of modulation of the PMN and N400 consistently differed between cohort and rhyme mismatches, with rhyme mismatches resulting in increased PMN responses compared to matching words that were not observed for cohort mismatches, and cohort mismatches resulting in larger Late N400 effects compared to matching words than those incurred by rhyme mismatches. These patterns were interpreted as evidence of incremental processing, or continuous mapping between presented input and expected word forms, as opposed to processing on the basis of global similarity.

However, although the Malins and Joanisse (2012) participants were English language learners living in an English-speaking environment at the time of testing, we only examined Mandarin language processing in these individuals. Hence, the cross-modal picture-spoken word matching paradigm has yet to be applied to study the modulatory effect of phonological cues on spoken word processing in two languages in the same set of participants. By performing this investigation, we may gain insights into the extent to which spoken word recognition in L2 English relies upon similar neurocognitive processes as those used to process L1 Mandarin in the same set of bilingual speakers. These findings could have implications for theories and models of bilingual spoken word recognition and second language learning, especially for language pairs that are typologically dissimilar (Kroll and Tokowicz, 2005; Van Heuven and Dijkstra, 2010).

### 1.3. *The present study*

Although prior studies have used ERPs to investigate the time course of spoken word processing in different languages (e.g., Desroches et al., 2009; Malins et al., 2014; Malins and Joanisse, 2012), much less is known about the neurocognitive processes that allow bilingual listeners to resolve phonological competition within multiple language systems, especially those that are typologically distinct (Mandarin vs. English). Accordingly, in the present study, we administered the same cross-modal picture-spoken word matching task in a group of Mandarin-English speakers using a within-subjects design to compare the timing and nature of the neurocognitive processes underlying phonological competition (cohort and rhyme) effects in L1 Mandarin and L2 English. When conducting this study, efforts were made to avoid cross-linguistic interference in lexical activation as much as possible (Wu and Thierry, 2010); these included an intervening non-language task between the Mandarin and English tasks as well as the exclusive use of each respective language for all task instructions (refer to section 2.2 of the Methods).

This work was motivated by several findings. As reviewed, there is debate as to whether the components of a syllable (onsets vs. rimes) might be weighted differently during Mandarin versus English word recognition in adult speakers (Malins and Joanisse, 2012; Zhao et al., 2011). In addition, previous work has shown that behaviorally, phonological awareness skills are moderately correlated between languages in Mandarin-English bilingual children for some measures but not others (partial correlation coefficients ranging from 0.08 to 0.51; Marinova-Todd et al., 2010). These differences at the behavioral level may result in differences in the neurocognitive processes engaged when resolving competiton amongst phonologically similar words in each language. Furthermore, recent work has shown that even in adult second language learners, listeners may display language-specific processing biases for different components of the syllable when processing Mandarin compared to English (e.g., vowels in Mandarin versus consonants in English; Weiner, 2019); these may also be reflected in brain potentials. Finally, some studies have reported that during L2 processing, various ERP components such as the N400 and P600 are delayed in peak latency, reduced in amplitude (Hahne, 2001; Kotz, 2009; Xue et al., 2013), or absent especially for beginning learners (Chen et al., 2007). Accordingly, it is intriguing to evaluate whether there are potential differences in the temporal dynamics of spoken word processing between L1 Mandarin and L2 English.

Although these previous studies suggest that the neurocognitive processes supporting spoken word recognition in L1 Mandarin may differ from those used to support spoken word recognition in L2 English, based on our previous work (Malins and Joanisse, 2012; Malins et al., 2014), we hypothesized that Mandarin-English bilinguals process spoken words incrementally in both languages using similar neurocognitive processes. Accordingly, in both languages, we expected to observe patterns of responses to word mismatches that are differentiable depending on the timing of the divergence between an expected word and a mismatch (i.e., word-initial versus word-final overlap). More specifically, we expected that in both languages, (1) rhyme mismatches would modulate the PMN compared to matching words, whereas cohort mismatches would not; (2) cohort mismatches would incur Late N400 effects compared to matching words that would be larger in magnitude than those incurred by rhyme mismatches.

## 2. Methods

### 2.1. *Participants*

We recruited forty Mandarin-English young adult bilingual participants drawn from the community of students at the University of Science and Technology Beijing in China (Mean age = 20.55, SD = 1.39; 24 female and 16 male). To assess Mandarin speakers with a range of English proficiency, participants were recruited from English majors and non-English majors. All participants were native speakers of Mandarin Chinese, were exposed to Mandarin outside of class hours, and used Mandarin for daily communication. When recruiting participants, exclusionary criteria included vision or hearing impairments, left handedness, a reported history of neurological disorders, or a reported history of language or learning disabilities.

To measure language proficiency in English, participants completed the following reading and language assessments: the Letter Word Identification subtest of the Woodcock-Johnson III Tests of Achievement (Woodcock et al., 2001), which was used to measure single word reading; subtests of the Gates-MacGinitie Reading Test Second Canadian Edition (Form 4, Level F) (MacGinitie et al., 2006), which were used to measure vocabulary and reading comprehension; and the Comprehensive Test of Phonological Processing (Wagner et al., 1999), which was used to measure phonemic awareness. Across the 34 participants included in the analysis of ERP data (after data from six participants were excluded for data quality purposes; see below), English proficiency was normally distributed based on the Shapiro-Wilk normality test ($W = .978$, $p = .716$), with proficiency defined as a composite of single word reading, vocabulary, and reading comprehension.

To verify that all participants had native Mandarin proficiency, participants also completed three tests of Mandarin proficiency: a single character reading test, in which participants read single characters that were matched in terms of character frequency, number of strokes, and the percentage rate of phonograms with Chinese characters taught in schools (Shu et al., 2003); a vocabulary test adapted to Mandarin from the Stanford-Binet Intelligence Scale vocabulary sub-test (Thorndike et al., 1986), in which participants were required to orally define 32 Mandarin two-character words; and a reading comprehension test adapted to Mandarin from items in the Gates-MacGinitie Reading Test Form 3 Level F, in which readers were presented with a series of 48 short passages, each of which was followed by a four-option multiple choice question.

To test nonverbal IQ, participants completed the Matrix Analogy Reasoning Set 2 and Set 4 (Naglieri, 1985). In addition to these assessments, participants were also asked to complete the Mandarin Chinese version of the Language History and Experience Questionnaire (LEAP-Q; Marian et al., 2007) as well as a short demographic questionnaire that contained several items related to medical history. Data from six participants were removed for data quality purposes during the preprocessing of ERP data: two who did not have usable data for both tasks, and four who did not have usable data for the English task (less than

50% of accepted trials after preprocessing and removal of incorrect trials). For the remaining 34 participants (20 female and 14 male), mean age was 20.59 (SD = 1.29). For additional demographic characteristics and assessment scores, refer to Table 1.

### 2.2. Stimuli and procedures

Stimuli in the ERP experiment were monosyllabic, highly imageable concrete nouns. Following Desroches et al. (2009), English and Mandarin stimulus lists were constructed separately in each language by deriving all mismatch types from a limited set of target items (16 in each language) that served as common references across stimulus types within each language (e.g., *bell* in English and *tang2* in Mandarin in the following examples). These mismatches were of the following three types: cohort, rhyme, or unrelated. Cohort mismatches shared onset information (and tone in Mandarin) with targets (e.g., BELL-*bed*; TANG2-*tao2*), rhyme mismatches shared word-final information (and tone) with targets (e.g., BELL-*shell*; TANG2-*lang2*), and unrelated items did not share any phonological overlap with targets (e.g., BELL-*ham*; TANG2-*xia1*). Mandarin stimuli were taken from the Malins et al. (2014) study; accordingly, the Mandarin task also included two additional mismatch types related to lexical tone (words sharing all phonemes but not tone with picture names, and words sharing only tone but not any phonemes with picture names) not included in the present analysis given our focus on between-language effects. To prevent inter-language interference or facilitation (Marian & Spivey, 2003a, 2003b), we verified that Mandarin translations of English stimuli, and vice-versa, did not overlap phonologically in onset or rime.

English stimulus sets (target-cohort-rhyme-unrelated) were balanced for frequency (Brysbaert and New, 2009) [$F(3,45) = .959$, $p = .41$], logarithmic frequency [$F(3,45) = .416$, $p = .73$], phonemic length [$F(3,45) = 1.63$, $p = .22$], and number of phonological neighbors (Balota et al., 2007) [$F(3,45) = 2.36$, $p = .11$]. Mandarin stimulus sets were balanced for frequency [$F(3,45) = .291$, $p = .83$], logarithmic frequency [$F(3,45) = .085$, $p = .97$], and phonemic length [$F(3,45) = 2.02$, $p = .13$], with frequency counts taken from the Modern Chinese Frequency Dictionary (1986).

Picture stimuli were color photos of objects on a white background resized to 275 x 275 pixels. For the English task, auditory stimuli were obtained from a repository of digital recordings at Haskins Laboratories. The speaker was an adult male native speaker of English who is a trained phonetician. For the Mandarin task, stimuli were recorded at the University of Western Ontario by an adult male native Mandarin speaker who is a trained news presenter. Both sets of stimuli were digitally recorded in a sound-proofed booth at a sampling rate of 44 kHz, with individual words spoken in isolation. Mandarin items had a mean duration of 437 ms, whereas English items had a mean duration of 389 ms. Within each language, duration was matched across stimulus sets (Mandarin: $F(3,45) = .173$, $p = .91$; English: $F(3,45) = 2.49$, $p = .07$).

Prior to each experimental task, participants completed a naming task in the respective language of testing (Mandarin or English) in which they were presented with pictures of items and asked to name them aloud. In cases in which they offered a name other than the intended name, they were told the intended name of the picture and asked to repeat it. Mean accuracy for the English and Mandarin naming tasks was 59% and 82%, respectively. Based on this data, the current sample of participants found it more difficult to orally produce appropriate picture names in English, consistent with English as their L2. However, when participants were told the intended names for items they did not name correctly, they indicated they were familiar with the intended names and agreed that they were well depicted by associated pictures. As noted further below, behavioral accuracy for the English matching task itself was high (89% collapsed across conditions) indicating that even though participants had difficulty overtly naming items in their L2, they were sufficiently familiar with the item names that they were able to successfully recognize whether or not the spoken words matched the pictures. To ensure comparability in the analysis of ERP effects across the two languages, only correct trials were included in all ERP analyses.

For the ERP trials, participants were seated 50 cm in front of a 24-in. CRT monitor. A fixation cross first appeared on screen for 250 ms, followed by a picture for 1500 ms. Next, while the picture remained on the screen, a spoken word stimulus was presented via a loudspeaker. Participants indicated via button press whether the spoken word matched or mismatched the picture. A blank screen was then presented for 1000 ms prior to the next trial. Participants were instructed not to blink during presentation of the stimuli, but rather to wait until presentation of the blank screen or fixation cross.

A practice block was presented prior to the task in each language, and consisted of six trials containing items not used in the actual experiment. For the English experimental task, there were 192 trials (96 match trials and 96 mismatch trials) presented in a pseudo-random order and divided into four equally sized blocks (48 trials each) with short rests between each set of trials. Mismatches consisted of 32 trials each of cohort, rhyme, and unrelated mismatches (i.e., two trials for each of the 16 sets; for example for the cohort pair *bell-bed* in English, in one trial BED was the picture and *bell* was the sound, and in the other trial BELL was the picture and *bed* was the sound). Match trials consisted of six trials for each of the 16 sets (e.g., for the set *bell-bed-shell-ham*, *bell* was the target in three trials, whereas *bed*, *shell*, and *ham* were the target for one trial each). In this way, each time a specific picture appeared on the screen, it was equally likely that the ensuing spoken word would either match or mismatch it. Mandarin trials consisted of 320 trials total (80 in each block), divided into 160 match trials and 160 mismatch trials (32 trials for each of the five mismatch types). In the Mandarin task, match trials consisted of ten trials for each of the 16 sets (e.g., for the set *tang2-tao2-lang2-tang1-niu2-xia1*, *tang2* was the target in five trials, whereas *tao2*, *lang2*, *tang1*, *niu2*, and *xia1* were the target for one trial each).

Order of administration of the English and Mandarin tasks was counterbalanced across participants. To minimize cross-linguistic

**Table 1**
Demographic characteristics of the 34 Mandarin-English bilingual participants (20 female) included in the analysis of ERP data.

| Measure (Maximum Possible score) | Means of Assessment | Mean | SD | Range |
|---|---|---|---|---|
| Age | – | 20.59 | 1.29 | 18-25 |
| Performance IQ (32) | Matrix Analogy Reasoning | 27.35 | 3.88 | 16-32 |
| Mandarin single character reading (150) | Chinese Single Character Identification | 144.21 | 2.79 | 135-148 |
| Mandarin reading comprehension (48) | Gates-MacGinitie Reading Test Form 3 Level F (translated into Mandarin) | 33.88 | 6.79 | 20-45 |
| Age of acquisition of Mandarin – speaking | LEAP-Q (Mandarin version) | 1.55 | 0.72 | 0.50-5 |
| Age of acquisition of Mandarin – reading | LEAP-Q (Mandarin version) | 4.87 | 1.77 | 1-8 |
| English vocabulary (65) | Gates-MacGinitie Reading Test Form 4 Level F | 24.76 | 12.39 | 7-57 |
| English single word reading (76) | Woodcock-Johnson III Tests of Achievement Letter Word Identification | 63.00 | 6.16 | 48-75 |
| English reading comprehension (48) | Gates-MacGinitie Reading Test Form 4 Level F | 20.91 | 9.38 | 3-44 |
| Age of acquisition of English - speaking | LEAP-Q (Mandarin version) | 7.04 | 3.37 | 1-14 |
| Age of acquisition of English - reading | LEAP-Q (Mandarin version) | 9.66 | 3.12 | 3-15 |

interference, an intervening non-language task (the AX-CPT cognitive control task) was conducted between the Mandarin and English tasks and took 45 min to administer (instructions for the AX-CPT were given in Mandarin). Furthermore, separate testers administered the English and Mandarin tasks, and all task instructions were given in the language in which each respective task was conducted. The two testers were both native Mandarin speakers and post-graduate students majoring in English language studies in the university. Both passed the national English Test for English majors, Band 8 (i.e., TEM 8).

All materials and procedures were approved by the Institutional Review Board at the School of Foreign Studies, University of Science and Technology Beijing. The participants gave their informed consent to participate in this study. All received paid compensation for participating.

### 2.3. Acquisition of ERP data

EEG data were acquired in a quiet room over a 1.5 h session. Continuous EEG data were recorded from 64 tin electrodes mounted in an elastic cap (Quik-Cap 64), positioned according to the International 10–20 system. During recording, all electrodes were referenced to the vertex (REF) electrode, with the GND electrode serving as the ground. Vertical and horizontal electrooculograms were recorded from bipolar pairs of vertical (VEOG) and lateral (HEOG) electrodes respectively placed above and below the left eye and the outer canthus of each eye. Electrode impedances were kept below 10 kΩ. Electrical signals were amplified with a Neuroscan Synamps 2 amplifier (60 Hz notch filter), using a band-pass filter of 0.1–100 Hz and a sampling frequency of 1000 Hz.

### 2.4. Analysis of ERP data

Data were processed offline in Curry 7 (Compumedics Limited). EEG data were re-referenced offline to the mean of the left and right mastoids mathematically. Trials were segmented into epochs spanning −100 to 800 ms relative to spoken word onset, and baseline corrected to the pre-stimulus period (i.e., when the fixation cross was presented). Trials containing eye-blinks, movement artifacts, or peak-to-peak deflections over 75 μV were rejected by an automatic procedure. Trials were also removed from analysis if behavioral responses were incorrect. Average ERP waveforms were filtered with a 0.1–20 Hz 24 dB bandpass filter for display and statistical analysis.

For the 34 participants with useable data (for details, see 2.1 Participants section), the average number of accepted trials in each experimental condition (after artifact rejection and exclusion of incorrect responses) was as follows: for the Mandarin task, 143/160 (89%) for the match condition, 28/32 (88%) for the cohort condition, 31/32 (97%) for the rhyme condition, and 31/32 (97%) for the unrelated condition; for the English task, 69/96 (72%) for the match condition, 24/32 (75%) for the cohort condition, 24/32 (75%) for the rhyme condition, and 27/32 (84%) for the unrelated condition.

Fifteen electrodes were selected for analysis and were divided into five columns – left lateral (F7, T7, P7), left medial (F3, C3, P3), midline (Fz, Cz, Pz), right medial (F4, C4, P4), right lateral (F8, T8, P8) – each with three levels for the factor "region" (frontal, central, and posterior). Statistical analyses were conducted by performing separate linear mixed-effects models for each component in each electrode column (optimization performed using Bound Optimization by Quadratic Approximation; BOBYQA; Powell, 2009). Each analysis tested for main effects and interactions between the following within-subjects factors: stimulus type (4: match, cohort, rhyme, unrelated), region (3: frontal, central, and posterior), and for the lateral and medial columns, hemisphere (2; left and right). Order of administration of the Mandarin and English tasks (i.e., Mandarin-first or English-first) was also included as an additional fixed effect of non-interest. Random effects terms included random intercepts for participants and random intercepts for stimulus type, region, and hemisphere within participants; these random intercepts for stimulus type, region, and hemisphere within participants were removed for the N400 analysis for the lateral column for the Mandarin task and the Late N400 analysis for the lateral column for the English task because their inclusion gave rise to a singular fit (Barr et al., 2013). For all models, we followed up on significant main effects and interactions for the stimulus type factor using Tukey-corrected pairwise tests.

Linear mixed-effects models were performed in R version 3.6.1 (R Core Team, 2019), using version 1.1-21 of the package *lme4* (Bates et al., 2015). Post-hoc analyses were carried out using version 1.4.5 of the package *emmeans* (Lenth et al., 2020). Topographic maps were plotted using version 0.4.1 of the package *erpR* (Arcara and Petrova, 2019). When visualizing waveforms, data were resampled to 100 Hz.

## 3. Results

### 3.1. Behavioral data

For both the Mandarin and English tasks, Fig. 1 displays mean error rates across stimulus types for the button press response as well as mean reaction times for correct trials, relative to word onset. Trials with reaction times less than 150 ms were considered invalid anticipations. A repeated measures ANOVA including within-subjects factors of language (2) and stimulus type (4) revealed an interaction between these two factors for error rates [$F(3,99) = 7.00$, $p < .001$, $\eta^2_G = .03$]. Simple main effects analysis revealed a main effect of stimulus type in Mandarin [$F(3,99) = 12.22$, $p < .0001$, $\eta^2_G = .10$] and English [$F(3,99) = 25.47$, $p < .0001$, $\eta^2_G = .19$], with Bonferroni-corrected post-hoc *t*-tests revealing that error rates were higher for cohort mismatches compared to each of the other stimulus types (Mandarin: cohort vs. match $p = .007$; cohort vs. rhyme $p = .004$; cohort vs. unrelated $p < .0001$; English: cohort vs. match $p = .01$; cohort vs. rhyme $p < .0001$; cohort vs. unrelated $p < .0001$); also in English we observed higher error rates for rhyme mismatches compared to unrelated mismatches ($p < .0001$).

For reaction times (RTs), there was a main effect of language [$F(3,99) = 28.74$, $p < .0001$, $\eta^2_G = .12$], marked by longer RTs for English compared to Mandarin, as well as a main effect of stimulus type [$F(3,99) = 67.70$, $p < .0001$, $\eta^2_G = .10$]. Bonferroni-corrected post-hoc *t*-tests (collapsing across language) revealed that RTs were longer for cohort mismatches compared to the other three stimulus types ($p < .0001$ for all three comparisons), longer for rhyme mismatches compared to match trials ($p < .0001$), longer for unrelated mismatches compared to match trials ($p = .015$), and longer for rhyme mismatches compared to unrelated mismatches ($p = .002$).

### 3.2. ERP data

Guided by Desroches et al. (2009), three latency windows were selected for statistical analyses: the PMN (230–310 ms), the N400 (310–410 ms), and the Late N400 (410–600 ms). ERP waveforms are shown for the Mandarin task in Fig. 2 and the English task in Fig. 3. Topographic maps for each of the three component windows are shown for the Mandarin task in Fig. 4 and the English task in Fig. 5. Full results from the linear mixed-effects models are reported for the Mandarin task in Table 2 and the English task in Table 3, whereas results for post-hoc contrasts between stimulus types are presented for the Mandarin task in Table 4 and the English task in Table 5.

#### 3.2.1. Mandarin task

*3.2.1.1. PMN (230–310 ms).* In the PMN analysis window, there was a significant main effect of stimulus type in the midline and medial columns as well as an interaction between stimulus type and hemisphere in the lateral column. Tukey-corrected post-hoc tests revealed that a

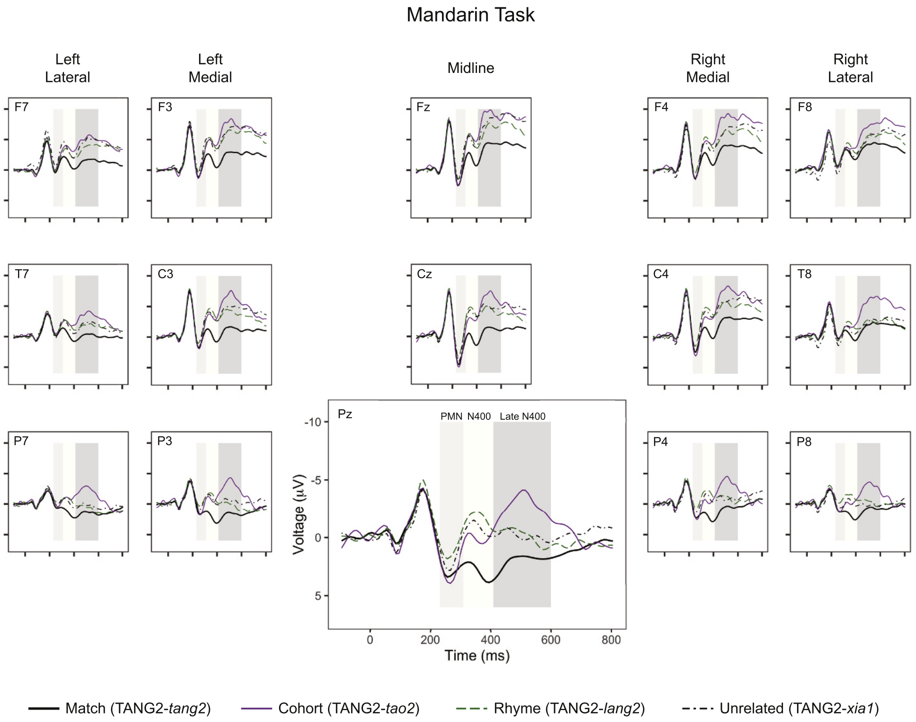


**Fig. 2.** Waveforms for the match, cohort, rhyme, and unrelated conditions in the Mandarin task. Boxes delineate the PMN, N400, and Late N400 windows that were selected for statistical analysis.

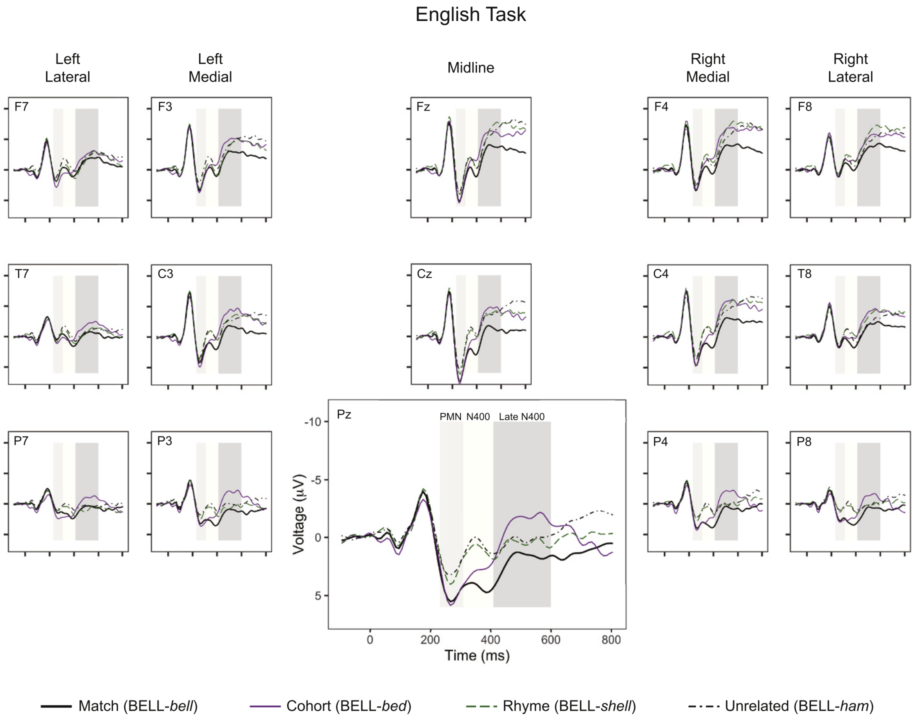


**Fig. 3.** Waveforms for the match, cohort, rhyme, and unrelated conditions in the English task. Boxes delineate the PMN, N400, and Late N400 windows that were selected for statistical analysis.

significantly larger PMN amplitude was observed in the rhyme mismatch condition compared to the match condition, and a marginally larger PMN amplitude was observed in the unrelated mismatch condition compared to the match condition. However, the cohort mismatch condition and the match condition did not differ from each other in PMN amplitude. In addition, when the different mismatch types

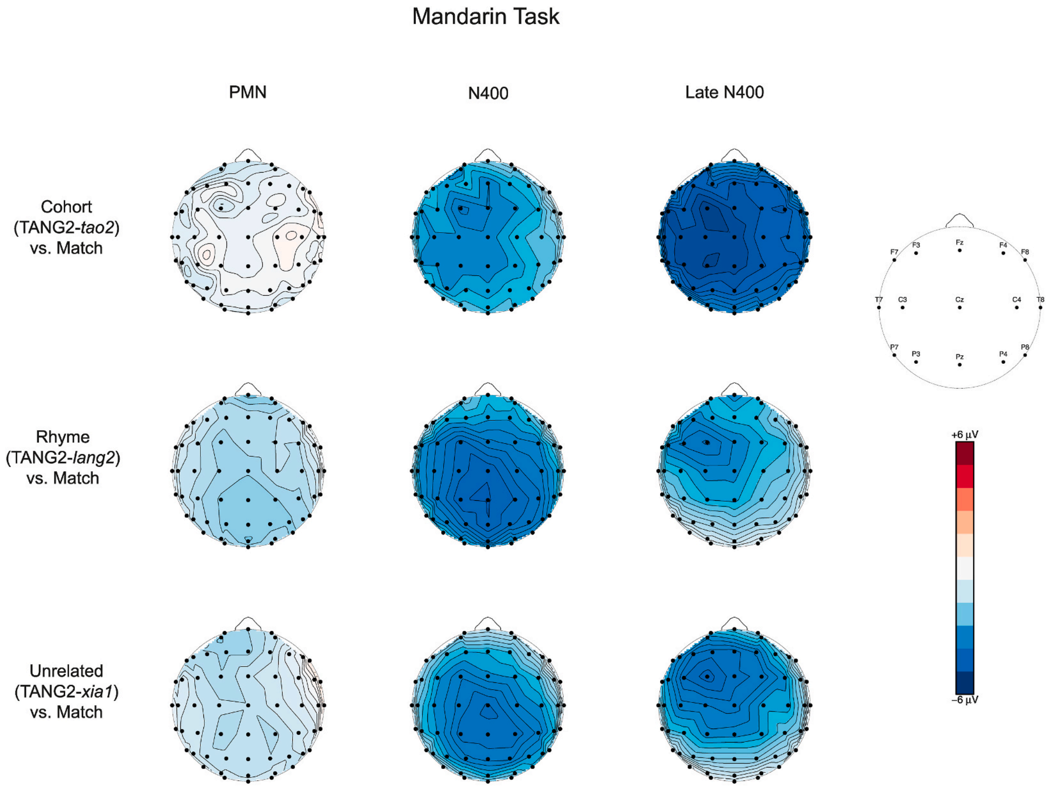


**Fig. 4.** Topographic maps for the differences in amplitude between the cohort, rhyme, and unrelated conditions respectively compared to the match condition in the Mandarin task.

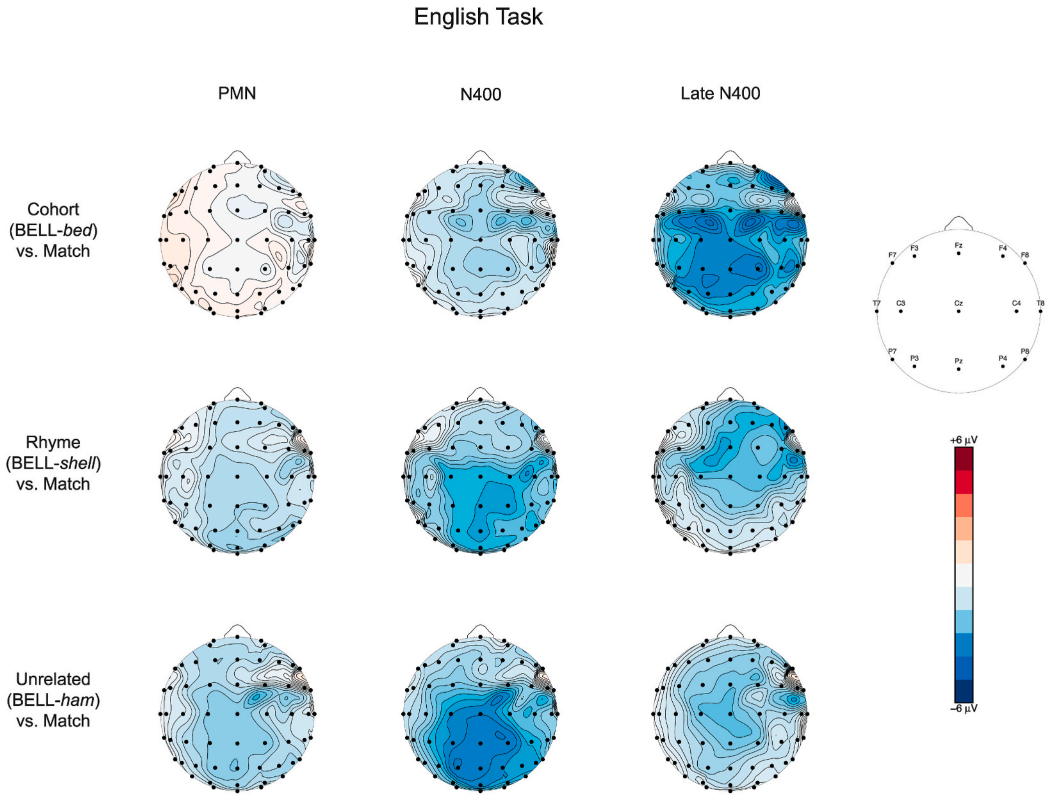


**Fig. 5.** Topographic maps for the differences in amplitude between the cohort, rhyme, and unrelated conditions respectively compared to the match condition in the English task.

**Table 2**
Summary of linear mixed-effects models for mean amplitude of the PMN, N400, and Late N400 for the Mandarin picture word-matching task. Models included a within-subjects factors of stimulus type (match, cohort mismatch, rhyme mismatch, and unrelated mismatch) and anterior-posterior region (frontal, central, posterior) for the midline column, and an additional within-subjects factor of hemisphere (left and right) for the medial and lateral columns. Models also included a between-subjects factor of order of administration of the two matching tasks (Mandarin-first or English-first).

| Effect | df | PMN (230–310 ms) | | | N400 (310–410 ms) | | | Late N400 (410–600 ms) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Midline | Medial | Lateral | Midline | Medial | Lateral | Midline | Medial | Lateral |
| Order of Administration *F* | 1,32 | .003 | .004 | .478 | .690 | .162 | .500 | .013 | .402 | 2.18 |
| *p* | | .960 | .952 | .495 | .412 | .690 | .485 | .910 | .530 | .150 |
| Stimulus Type *F* | 3,99 | 4.44 | 3.49 | .697 | 12.21 | 10.65 | 14.88 | 16.55 | 17.77 | 16.35 |
| *p* | | .006 | .018 | .556 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |
| Region *F* | 2,66 | 15.20 | 10.18 | 16.09 | 53.25 | 63.50 | 47.42 | 84.70 | 93.87 | 71.79 |
| *p* | | <.0001 | <.001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |
| Hemisphere *F* | 1,33 | – | 4.12 | .746 | – | .172 | .086 | – | 5.85 | 3.73 |
| *p* | | – | .051 | .394 | – | .681 | .770 | – | .021 | .062 |
| Region × Stim Type *F* | 6,198 (6,561) | 1.73 | .688 | 1.17 | 6.06 | 1.98 | .492 | 4.25 | 2.66 | 1.08 |
| *p* | | .116 | .659 | .322 | <.0001 | .066 | .815 | <.0001 | .015 | .373 |
| Hemisphere × Stim Type *F* | 3,561 | – | 2.05 | 2.84 | – | 1.24 | .729 | – | 1.31 | 1.92 |
| *p* | | – | .106 | .037 | – | .294 | .535 | – | .270 | .126 |
| Hemisphere × Region *F* | 2,561 | – | .216 | 2.47 | – | .433 | 1.02 | – | 2.40 | 2.30 |
| *p* | | – | .805 | .085 | – | .649 | .362 | – | .092 | .101 |
| Hemi × Region × Stim Type *F* | 6,561 | – | .692 | .787 | – | .496 | .303 | – | .894 | .593 |
| *p* | | – | .656 | .581 | – | .811 | .935 | – | .499 | .736 |

**Table 3**
Summary of linear mixed-effects models for mean amplitude of the PMN, N400, and Late N400 for the English picture word-matching task. Models included a within-subjects factors of stimulus type (match, cohort mismatch, rhyme mismatch, and unrelated mismatch) and anterior-posterior region (frontal, central, posterior) for the midline column, and an additional within-subjects factor of hemisphere (left and right) for the medial and lateral columns. Models also included a between-subjects factor of order of administration of the two matching tasks (Mandarin-first or English-first).

| Effect | df | PMN (230–310 ms) | | | N400 (310–410 ms) | | | Late N400 (410–600 ms) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Midline | Medial | Lateral | Midline | Medial | Lateral | Midline | Medial | Lateral |
| Order of Administration *F* | 1,32 | .421 | .027 | .107 | 1.05 | 1.06 | .308 | .015 | .131 | .003 |
| *p* | | .521 | .870 | .746 | .314 | .312 | .583 | .902 | .719 | .957 |
| Stimulus Type *F* | 3,99 | 5.29 | 3.27 | 2.49 | 8.91 | 6.83 | 5.38 | 9.98 | 9.99 | 5.38 |
| *p* | | .002 | .024 | .064 | <.0001 | <.001 | .002 | <.0001 | <.0001 | .001 |
| Region *F* | 2,66 | 11.90 | 4.91 | 7.77 | 32.01 | 35.51 | 31.37 | 63.64 | 67.51 | 47.72 |
| *p* | | <.0001 | .010 | <.001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |
| Hemisphere *F* | 1,33 | – | .019 | .016 | – | 1.96 | .580 | – | 5.21 | 27.78 |
| *p* | | – | .890 | .901 | – | .171 | .452 | – | .029 | <.0001 |
| Region × Stim Type *F* | 6,198 (6,561) | .813 | 1.41 | 1.46 | 7.39 | 4.78 | 1.29 | 3.82 | 2.26 | .438 |
| *p* | | .561 | .210 | .188 | <.0001 | <.0001 | .262 | .001 | .037 | .853 |
| Hemisphere × Stim Type *F* | 3,561 | – | 2.42 | 2.15 | – | 3.21 | 3.11 | – | 3.29 | 1.28 |
| *P* | | – | .066 | .094 | – | .023 | .026 | – | .020 | .279 |
| Hemisphere × Region *F* | 2,561 | – | 4.21 | 11.62 | – | 5.53 | 8.78 | – | 13.79 | 6.97 |
| *P* | | – | .015 | <.0001 | – | .004 | <.001 | – | <.0001 | <.001 |
| Hemi × Region × Stim Type *F* | 6,561 | – | .697 | .758 | – | .832 | .512 | – | .755 | .087 |
| *P* | | – | .652 | .603 | – | .545 | .795 | – | .606 | .998 |

were compared to each other, the PMN was larger for rhyme mismatches compared to cohort mismatches.

*3.2.1.2. N400 (310–410 ms).* In the N400 analysis window, there was a significant interaction between stimulus type and electrode region in the midline column and a main effect of stimulus type in the medial and lateral columns. Tukey-corrected post-hoc tests revealed that, compared to the match condition, larger N400 amplitudes were observed in all three mismatch conditions. However, the three mismatch conditions did not differ from each other in N400 amplitude.

*3.2.1.3. Late N400 (410–600 ms).* In the Late N400 analysis window, there were significant interactions between stimulus type and electrode region in the midline and medial columns as well as a main effect of stimulus type in the lateral column. Tukey-corrected post-hoc tests revealed that the Late N400 was larger for cohort, rhyme, and unrelated mismatches compared to the match condition. In addition, when the different mismatch types were compared to each other, the Late N400 was larger for cohort mismatches compared to rhyme and unrelated mismatches.

### 3.2.2. *English Task*

*3.2.2.1. PMN (230–310 ms).* In the PMN analysis window, there was a significant main effect of stimulus type in the midline and medial columns. Tukey-corrected post-hoc tests revealed that a significantly larger PMN amplitude was observed in the unrelated mismatch condition compared to the match condition, and a marginally larger PMN amplitude was observed in the rhyme mismatch condition compared to the match condition. However, the cohort mismatch condition and the match condition did not differ from each other in PMN amplitude. In addition, when the different mismatch types were compared to each other, the PMN was significantly larger for unrelated mismatches compared to cohort mismatches and marginally larger for rhyme mismatches compared to cohort mismatches.

*3.2.2.2. N400 (310–410 ms).* In the N400 analysis window, there was a significant interaction between stimulus type and electrode region in the midline and medial columns as well as a significant interaction between stimulus type and hemisphere in the lateral column. Tukey-corrected post-hoc tests revealed that larger N400 amplitudes were observed in

**Table 4**
Summary of Tukey-corrected post-hoc contrasts for the Mandarin task.

| | PMN | | | N400 | | | Late N400 | | |
|---|---|---|---|---|---|---|---|---|---|
| Contrast | Electrode Location | *t* | *p* | Electrode Location | *t* | *p* | Electrode Location | *t* | *p* |
| Cohort vs. Match | midline column | -.317 | .989 | Fz | −4.22 | .003 | Fz | −6.43 | <.0001 |
| | | | | Cz | −4.04 | .005 | Cz | −6.86 | <.0001 |
| | | | | Pz | −3.71 | .016 | Pz | −6.28 | <.0001 |
| | medial column | -.459 | .968 | medial column | −4.18 | <.001 | medial frontal | −6.78 | <.0001 |
| | | | | | | | medial central | −6.93 | <.0001 |
| | | | | | | | medial posterior | −6.52 | <.0001 |
| | left lateral column | −1.35 | .878 | lateral column | −5.75 | <.0001 | lateral column | −6.91 | <.0001 |
| | right lateral column | -.026 | .999 | | | | | | |
| Rhyme vs. Match | midline column | −2.97 | .019 | Fz | −3.95 | .007 | Fz | −3.61 | .022 |
| | | | | Cz | −5.85 | <.0001 | Cz | −4.24 | .003 |
| | | | | Pz | −5.95 | <.0001 | Pz | −2.83 | .181 |
| | medial column | −2.84 | .028 | medial column | −5.15 | <.0001 | medial frontal | −3.97 | .006 |
| | | | | | | | medial central | −3.74 | .014 |
| | | | | | | | medial posterior | −2.45 | .381 |
| | left lateral column | −1.32 | .892 | lateral column | −5.81 | <.0001 | lateral column | −2.49 | .067 |
| | right lateral column | −1.13 | .950 | | | | | | |
| Unrelated vs. Match | midline column | −2.43 | .078 | Fz | −3.61 | .022 | Fz | −5.18 | <.001 |
| | | | | Cz | −5.25 | <.0001 | Cz | −4.89 | <.001 |
| | | | | Pz | −4.77 | <.001 | Pz | −2.81 | .187 |
| | medial column | −2.00 | .195 | medial column | −4.26 | <.001 | medial frontal | −4.70 | <.001 |
| | | | | | | | medial central | −4.69 | <.001 |
| | | | | | | | medial posterior | −2.82 | .185 |
| | left lateral column | −1.99 | .491 | lateral column | −4.11 | <.001 | lateral column | −2.83 | .028 |
| | right lateral column | .764 | .995 | | | | | | |
| Cohort vs. Rhyme | midline column | 2.65 | .045 | Fz | -.278 | .999 | Fz | −2.82 | .184 |
| | | | | Cz | 1.81 | .808 | Cz | −2.62 | .278 |
| | | | | Pz | 2.23 | .531 | Pz | −3.45 | .035 |
| | medial column | 2.38 | .089 | medial column | .977 | .763 | medial frontal | −2.81 | .188 |
| | | | | | | | medial central | −3.19 | .073 |
| | | | | | | | medial posterior | −4.71 | .004 |
| | left lateral column | -.036 | .999 | lateral column | .064 | .999 | lateral column | −4.41 | <.001 |
| | right lateral column | 1.10 | .956 | | | | | | |
| Cohort vs. Unrelated | midline column | 2.11 | .157 | Fz | -.619 | .999 | Fz | −1.25 | .984 |
| | | | | Cz | 1.21 | .988 | Cz | −1.97 | .713 |
| | | | | Pz | 1.05 | .996 | Pz | −3.47 | .034 |
| | medial column | 1.54 | .417 | medial column | .083 | .999 | medial frontal | −2.08 | .635 |
| | | | | | | | medial central | −2.24 | .522 |
| | | | | | | | medial posterior | −3.71 | .016 |
| | left lateral column | .640 | .998 | lateral column | −1.64 | .356 | lateral column | −4.07 | <.001 |
| | right lateral column | -.790 | .994 | | | | | | |
| Rhyme vs. Unrelated | midline column | -.543 | .948 | Fz | -.340 | .999 | Fz | 1.57 | .916 |
| | | | | Cz | -.606 | .999 | Cz | .655 | .999 |
| | | | | Pz | −1.18 | .990 | Pz | -.014 | .999 |
| | medial column | -.831 | .838 | medial column | -.894 | .808 | medial frontal | .727 | .999 |
| | | | | | | | medial central | .950 | .998 |
| | | | | | | | medial posterior | .368 | .999 |
| | left lateral column | .676 | .998 | lateral column | −1.71 | .322 | lateral column | .339 | .987 |
| | right lateral column | −1.89 | .559 | | | | | | |

the rhyme and unrelated mismatch conditions compared to the match condition but not in the cohort mismatch condition compared to the match condition. In addition, when the different mismatch types were compared to each other, the N400 was larger for unrelated mismatches compared to cohort mismatches.

*3.2.2.3. Late N400 (410–600 ms).* In the Late N400 analysis window, there were significant interactions between stimulus type and electrode region in the midline and medial columns, a significant interaction between stimulus type and hemisphere in the medial column, and a main effect of stimulus type in the lateral column. Tukey-corrected post-hoc tests revealed that the Late N400 was larger for cohort, rhyme, and unrelated mismatches compared to the match condition. In addition, when the different mismatch types were compared to each other, the Late N400 was significantly larger for cohort mismatches compared to rhyme mismatches and marginally larger for cohort mismatches compared to unrelated mismatches.

## 4. Discussion

The present study addressed whether the neurocognitive processes that allow bilingual listeners to resolve phonological competition differ between first- and second-language processing when those languages differ widely in phonological and morphological structure. To satisfy this aim, we used a cross-modal picture-spoken word matching ERP paradigm to investigate the temporal dynamics of first language (L1) Mandarin and second language (L2) English processing in the same set of Mandarin-English adult bilingual speakers.

ERP responses were evaluated for expected versus unexpected auditory words. The modulatory effect on ERP responses was determined by the extent to which the listener detected mismatches between presented words compared to the words they expected to hear, accordingly reflecting sensitivity to phonological information and the ability to suppress lexical alternatives (Malins and Joanisse, 2012; McMurray et al., 2010). More specifically, the present study examined the interference effects of different types of mismatches (i.e., cohorts, rhymes, and unrelated words) on ERP components associated with phonological and lexical processing (i.e., the PMN, N400, and Late

**Table 5**
Summary of Tukey-corrected post-hoc contrasts for the English task.

| | PMN | | | N400 | | | Late N400 | | |
|---|---|---|---|---|---|---|---|---|---|
| Contrast | Electrode Location | *t* | *p* | Electrode Location | *t* | *p* | Electrode Location | *t* | *p* |
| Cohort vs. Match | midline column | -.001 | .999 | Fz | −2.34 | .454 | Fz | −4.12 | .004 |
| | | | | Cz | −2.35 | .446 | Cz | −5.18 | <.0001 |
| | | | | Pz | −2.01 | .686 | Pz | −5.18 | <.0001 |
| | medial column | .248 | .995 | medial frontal | −2.18 | .569 | medial frontal | −4.52 | <.001 |
| | | | | medial central | −2.02 | .682 | medial central | −5.04 | <.001 |
| | | | | medial posterior | −1.57 | .918 | medial posterior | −5.05 | <.001 |
| | | | | | | | left medial column | −4.98 | <.001 |
| | | | | | | | right medial column | −5.29 | <.0001 |
| | | | | left lateral column | -.086 | .999 | lateral column | −3.96 | <.001 |
| | | | | right lateral column | −1.86 | .584 | | | |
| Rhyme vs. Match | midline column | −2.55 | .059 | Fz | −2.48 | .360 | Fz | −4.19 | .003 |
| | | | | Cz | −4.37 | .002 | Cz | −4.00 | .006 |
| | | | | Pz | −4.76 | <.001 | Pz | −2.26 | .510 |
| | medial column | −1.84 | .262 | medial frontal | −2.15 | .587 | medial frontal | −3.36 | .044 |
| | | | | medial central | −3.86 | .009 | medial central | −3.58 | .022 |
| | | | | medial posterior | −4.07 | .004 | medial posterior | −1.52 | .932 |
| | | | | | | | left medial column | −2.04 | .460 |
| | | | | | | | right medial column | −3.91 | .004 |
| | | | | left lateral column | −1.37 | .870 | lateral column | −2.02 | .181 |
| | | | | right lateral column | −3.99 | .003 | | | |
| Unrelated vs. Match | midline column | −3.05 | .016 | Fz | −2.49 | .354 | Fz | −3.17 | .078 |
| | | | | Cz | −5.13 | <.001 | Cz | −4.00 | .006 |
| | | | | Pz | −5.41 | <.0001 | Pz | −2.69 | .241 |
| | medial column | −2.29 | .108 | medial frontal | −2.52 | .335 | medial frontal | −2.63 | .274 |
| | | | | medial central | −3.84 | .010 | medial central | −2.64 | .265 |
| | | | | medial posterior | −4.88 | <.001 | medial posterior | −2.00 | .695 |
| | | | | | | | left medial column | −2.45 | .229 |
| | | | | | | | right medial column | −2.66 | .144 |
| | | | | left lateral column | −2.82 | .098 | lateral column | 1.44 | .472 |
| | | | | right lateral column | −3.31 | .025 | | | |
| Cohort vs. Rhyme | midline column | 2.55 | .059 | Fz | .142 | .999 | Fz | .070 | .999 |
| | | | | Cz | 2.02 | .681 | Cz | −1.18 | .990 |
| | | | | Pz | 2.75 | .216 | Pz | −2.92 | .146 |
| | medial column | 2.09 | .165 | medial frontal | -.025 | .999 | medial frontal | −1.15 | .992 |
| | | | | medial central | 1.85 | .789 | medial central | −1.46 | .949 |
| | | | | medial posterior | 2.50 | .347 | medial posterior | −3.53 | .027 |
| | | | | | | | left medial column | −2.94 | .073 |
| | | | | | | | right medial column | −1.37 | .868 |
| | | | | left lateral column | 1.28 | .904 | lateral column | −1.94 | .213 |
| | | | | right lateral column | 2.13 | .400 | | | |
| Related | midline column | 3.04 | .016 | Fz | .152 | .999 | Fz | -.947 | .999 |
| | | | | Cz | 2.78 | .201 | Cz | −1.18 | .990 |
| | | | | Pz | 3.40 | .041 | Pz | −2.49 | .358 |
| | medial column | 2.54 | .061 | medial frontal | .346 | .999 | medial frontal | −1.89 | .764 |
| | | | | medial central | 1.82 | .803 | medial central | −2.40 | .413 |
| | | | | medial posterior | 3.32 | .051 | medial posterior | −3.05 | .104 |
| | | | | | | | left medial column | −2.54 | .189 |
| | | | | | | | right medial column | −2.62 | .158 |
| | | | | left lateral column | 2.73 | .121 | lateral column | −2.52 | .058 |
| | | | | right lateral column | 1.45 | .831 | | | |
| Rhyme vs. Unrelated | midline column | .497 | .960 | Fz | .010 | .999 | Fz | −1.02 | .997 |
| | | | | Cz | .763 | .999 | Cz | .003 | .999 |
| | | | | Pz | .652 | .999 | Pz | .433 | .999 |
| | medial column | .450 | .969 | medial frontal | .371 | .999 | medial frontal | -.735 | .999 |
| | | | | medial central | -.025 | .999 | medial central | -.938 | .999 |
| | | | | medial posterior | .815 | .999 | medial posterior | .474 | .999 |
| | | | | | | | left medial column | .406 | .999 |
| | | | | | | | right medial column | −1.25 | .915 |
| | | | | left lateral column | 1.45 | .834 | lateral column | -.578 | .939 |
| | | | | right lateral column | -.678 | .998 | | | |

N400). Overall, the current sample of late bilinguals showed lower accuracy and longer reaction times in L2 English compared to L1 Mandarin, reflective of their lower proficiency in English. Yet, ERP responses in the present study revealed commonalities in the neurocognitive processes underlying Mandarin versus English spoken word recognition that were not entirely apparent in the observed patterns of behavioral responses.

### *4.1. Comparing the neurocognitive processes supporting L1 Mandarin and L2 English spoken word recognition*

To compare the timing and nature of L1 Mandarin and L2 English spoken word recognition, we evaluated ERP responses from the same set of adult Mandarin-English bilingual participants to the same type of mismatches (cohort, rhyme, unrelated) in each language. Specifically, in the case of rhyme mismatches (e.g., Mandarin: TANG2-*lang2*; English: BELL-*shell*), the picture-spoken word pair was different in the first phoneme, whereas in unrelated mismatches (e.g., Mandarin: TANG2-

*xia1*: English: BELL-*ham*), all phonemes – as well as tone in Mandarin – were different. In the case of cohort mismatches (e.g., Mandarin: TANG2-*tao2*; English: BELL-*bed*), there was a difference in word-final phonemes. As word initial phonemes in rhyme and unrelated mismatches differed from expectations, mismatches in these two conditions were signaled very early during the unfolding of spoken forms. Conversely, word-initial phonemes in the cohort condition matched expectations, so differences were not signaled until later in the spoken form. Therefore, effects for cohort versus rhyme mismatches are reflective of different stages of word recognition.

When comparing the different stages of Mandarin and English spoken word recognition, we uncovered evidence suggesting that processing patterns were differentiable according to mismatch types, yet similar across the two languages. First, as hypothesized, in both languages the PMN was modulated for rhyme mismatches compared to matching words, whereas this effect was not observed for cohort mismatches compared to matching words. This reflects sensitivity to initial phoneme mismatch in both languages (Connolly and Phillips, 1994; Newman et al., 2003). Second, as hypothesized, in both languages cohort mismatches induced larger Late N400 responses compared to matching words than rhyme mismatches, and furthermore, when the two mismatch conditions were compared directly to each other, the Late N400 was larger for cohort mismatches compared to rhyme mismatches. As the N400 and Late N400 index interference or competition from other lexical items (Van Petten, Coulson, Rubin, Plante and Parks, 1999), these results are reflective of an increased difficulty in processing for cohort mismatches, likely because of increased lexical competition based on bottom-up cues (Desroches et al., 2009). This pattern of ERP results is also complemented by the behavioral data, as higher error rates and longer reaction times were observed for cohort mismatches compared to the other mismatch conditions in both languages.

One difference that we observed between languages is that in Mandarin, the cohort mismatch condition showed differences from the match condition in both the N400 and Late N400 windows, whereas in English, the cohort mismatch condition was only different from the match condition in the Late N400 window. It is possible that this difference is indicative of increased difficulty in resolving bottom-up competition in L2 English compared to L1 Mandarin. However, another possible explanation is differences between languages in the point of acoustic divergence between cohort mismatches and expected words. To address this possibility, we calculated the point of divergence in formant frequency (F1 and F2) between cohort mismatches and expected words in both languages (e.g., *bell-bed*; *tang2-tao2*) using the program Praat version 6.1.10 (Boersma and Weenink, 2020). This analysis revealed that the point of divergence between cohort mismatches and targets was later in English than Mandarin in terms of F1 ($p < .028$) – although not different in terms of F2 ($p = .418$) – lending support to this explanation of differences between the Mandarin and English tasks in when cohort mismatches were signaled.

In addition to this difference between languages, we also noted a discrepancy from previous studies (e.g., Desroches et al., 2009; Malins et al., 2014) in that we did not observe any evidence for facilitated processing of word-final phonemes for rhyme mismatches in either language. Behaviorally, rhyme mismatches resulted in longer reaction times than unrelated mismatches when collapsing across language, perhaps reflecting increased difficulty in this condition resulting from competition on the basis of word-final overlap. This finding indicates sensitivity to word-final overlap in both languages. However, unlike prior studies which reported a reduction of the Late N400 for rhyme mismatches (English: Desroches et al., 2009; Mandarin: Malins et al., 2014), we did not observe this Late N400 reduction in either language. Among other factors, this discrepancy may be due to differences from previous studies in language background, language proficiency, and language use of the participant sample.

Nonetheless, in summary, our results show that spoken word recognition in both Mandarin and English is incremental and takes into account the temporal structure of words as they unfold (Liu et al., 2006; Huang et al., 2014). These data do not support the view that Mandarin spoken word recognition relies to a greater extent on global similarity in whole syllable structure compared to English (Zhao et al., 2011). Moreover, this finding of similarity in L1 and L2 processing, despite considerable differences in phonological and morphological structure between languages, implies commonality in the neurocognitive processes governing L1 and L2 spoken word recognition. This point is especially underscored because we observed this commonality in the current sample of exclusively unbalanced bilingual participants – with a strong language dominance for L1 Mandarin – who may have been less likely to exhibit this pattern of effects as opposed to more balanced bilinguals. With that said, future studies should discern how this commonality in neurocognitive processing is influenced by individual differences in factors such as age of second language acquisition, second language proficiency, and second language use.

### 4.2. *Implications for models of spoken word recognition*

These results have implications for models of bilingual spoken language processing (Kroll and Tokowicz, 2005; Van Heuven and Dijkstra, 2010). First, based on the current finding of incremental processing in both L1 Mandarin and L2 English, the results from this study support continuous mapping models, and could be captured in an interactive-activation framework such as that adopted in the TRACE model (McClelland and Elman, 1986), provided that relevant modifications are included for bilingual listeners (e.g., BIMOLA, the Bilingual Model of Lexical Access; Grosjean, 1997; BLINCS, the Bilingual Language Interaction Network for Comprehension of Speech; Shook and Marian, 2013). This interpretation stands in contrast to previous work suggesting that models that compute similarity on the basis of global overlap without respect to timing (such as the neighborhood activation model; Luce and Pisoni, 1998) are perhaps more appropriate for Mandarin than are continuous mapping models (Zhao et al., 2011).

Second, the current results suggest that bilingual listeners map speech input onto lexical representations in a similar fashion in their first and second languages, even in cases in which phonological features are considerably different between languages. Intriguingly, the current results may support bilingual models that incorporate a common layer for phonological representations in Mandarin and English (e.g., DevLex-II; Zhao and Li, 2010). However, as Kroll and Tokowicz (2005) point out, there is a risk of conflating processing with representation when considering these models.

Finally, although not examined in the current study, previous studies have shown that lexical tone – a critical feature of Mandarin Chinese – is an important contributor to lexical access in Mandarin-English bilingual listeners (Wang et al., 2017; Wang et al., 2020). Although we have previously modeled Mandarin lexical tone in an interactive activation framework (Shuai and Malins, 2017), we have only done this in a monolingual context. Therefore, future models of bilingual spoken word recognition should be developed using a framework that allows them to capture the processes involved in L1 and/or L2 tonal processing.

## 5. Conclusions

Using a cross-modal picture-spoken word matching ERP paradigm, we present evidence that the neurocognitive processes that allow bilingual individuals to resolve phonological competition in L1 Mandarin are similar to those that allow listeners to resolve phonological competition in L2 English. These findings support models of bilingual spoken word processing that adopt a similar interactive-activation framework for both L1 and L2 processing. Beyond this, the current findings offer support for the view that the brain exhibits remarkable commonalities in processing across languages irrespective of differences in phonological and morphological structure (Rueckl et al., 2015).

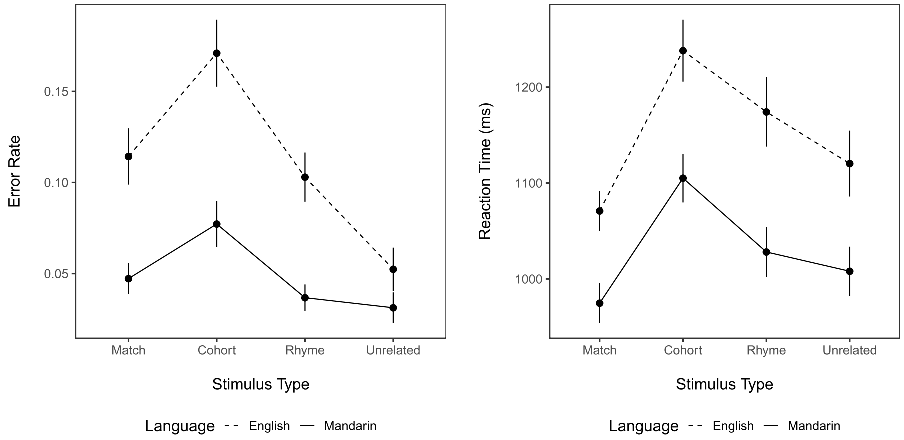


**Fig. 1.** Mean error rates and mean reaction time for correct trials for the button press response for the Mandarin and English picture-spoken word matching tasks. Error bars indicate the standard error of the mean.

## Declaration of competing interest

None.

## CRediT authorship contribution statement

**Jin Xue:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Banban Li:** Investigation. **Rong Yan:** Project administration. **Jeffrey R. Gruen:** Writing - review & editing, Supervision. **Tianli Feng:** Investigation. **Marc F. Joanisse:** Conceptualization, Methodology, Writing - review & editing. **Jeffrey G. Malins:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration.

## Acknowledgments

This work was supported by a grant from the Beijing Social Science Fund, China (19YYB008) to the first author. Support for JGM and JRG was provided by the Manton Foundation as well as NIH/Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) P50 HD027802. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

Arcara, G., Petrova, A., 2019. erpR: event-related potentials (ERP) analysis, graphics and utility functions. R package version 0.4.1/r116. https://R-Forge.R-project.org/projects/erpr/.

Archibald, L., Joanisse, M.F., 2011. Electrophysiological responses to coarticulatory and word level miscues. J. Exp. Psychol. Hum. Percept. Perform. 37 (4), 1275–1291.

Balota, D.A., Yap, M.J., Hutchison, K.A., Cortese, M.J., Kessler, B., Loftis, B., Treiman, R., 2007. The English lexicon project. Behav. Res. Methods 39 (3), 445–459.

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J. Mem. Lang. 68 (3), 255–278.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using *lme4*. Journal of Stastical Software 67 (1), 1–48.

Boersma, P., Weenink, D., 2020. Praat: doing phonetics by computer [Computer program]. retrieved 28 March 2020 from. http://www.praat.org/. Version 6.1.10.

Brysbaert, M., New, B., 2009. Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behav. Res. Methods 41 (4), 977–990.

Chen, J.Y., O'Séaghdha, P.G., Chen, T.M., 2016. The primacy of abstract syllables in Chinese word production. J. Experim. Psychol. Learn. Mem. Cogn. 42 (5), 825–836. https://doi.org/10.1037/a0039911.

Chen, L., Shu, H.U.A., Liu, Y., Zhao, J., Li, P., 2007. ERP signatures of subject-verb agreement in L2 learning. Bilingual. Lang.Cogn. 10 (2), 161–174. https://doi.org/10.1017/S136672890700291X.

Clahsen, H., Felser, C., 2006. How native-like is non-native language processing? Trends Cognit. Sci. 10 (12), 564–570. https://doi.org/10.1016/j.tics.2006.10.002.

Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J. Cognit. Neurosci. 6 (3), 256–266. https://doi.org/10.1162/jocn.1994.6.3.256.

Costa, A., Sebastián-Gallés, N., 2014. How does the bilingual experience sculpt the brain? Nat. Rev. Neurosci. 15 (5), 336–345. https://doi.org/10.1038/nrn3709.

Del Maschio, N., Abutalebi, J., 2019. Language organization in the bilingual and multilingual brain. In: Schwieter, J.W. (Ed.), The Handbook of the Neuroscience of Multilingualism. John Wiley & Sons Ltd, pp. 199–213. https://doi.org/10.1002/9781119387725.ch9.

Desroches, A.S., Newman, L., Robertson, E.K., Joanisse, M.F., 2013. Electrophysiological indices of phonological impairments in dyslexia. J. Speech Lang. Hear. Res. 56 (1), 250–265.

Desroches, A.S., Newman, R.L., Joanisse, M.F., 2009. Investigating the time course of spoken word recognition: electrophysiological evidence for the influences of phonological similarity. J. Cognit. Neurosci. 21 (10), 1893–1906. https://doi.org/10.1162/jocn.2008.21142.

Golestani, N., 2016. Neuroimaging of phonetic perception in bilinguals. Biling. Lang. Cognit. 19 (4), 674–682. https://doi.org/10.1017/S1366728915000644.

Grosjean, F., 1997. Processing mixed language: issues, findings, and models. In: De Groot, A.M.B., Kroll, J.F. (Eds.), Tutorials in Bilingualism: Psycholinguistic Perspectives. Psychology Press, New York, pp. 225–254.

Hahne, A., 2001. What's different in second-language processing? Evidence from event-related brain potentials. J. Psycholing. Res. 30 (3), 251–266.

Higby, E., Kim, J., Obler, L.K., 2013. Multilingualism and the brain. Annu. Rev. Appl. Ling. 33, 68–101. https://doi.org/10.1017/S0267190513000081.

Huang, X., Yang, J.C., Zhang, Q., Guo, C., 2014. The time course of spoken word recognition in Mandarin Chinese: A unimodal ERP study. Neuropsychologia 63, 165–174. https://doi.org/10.1016/j.neuropsychologia.2014.08.015.

Kornilov, S.A., Magnuson, J.S., Rakhlin, N., Landi, N., Grigorenko, E.L., 2015. Lexical processing deficits in children with developmental language disorder: an event-related potentials study. Dev. Psychopathol. 27 (2), 459–476. https://doi.org/10.1017/s0954579415000097.

Kotz, S.A., 2009. A critical review of ERP and fMRI evidence on L2 syntactic processing. Brain Lang. 109 (2–3), 68–74.

Kroll, J.F., Tokowicz, N., 2005. Models of bilingual representation and processing: looking back and to the future. In: Kroll, J.F., De Groot, A.M.B. (Eds.), Handbook of Bilingualism: Psycholinguistic Approaches. Oxford University Press, New York, pp. 531–553.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62 (1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Kutas, M., Hillyard, S., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207 (4427), 203–205.

Lenth, R., Singmann, H., Love, J., Buerkner, P., Herve, M., 2020. emmeans: estimated marginal means. R package version 1.4.4.

Liu, Y., Shu, H., Wei, J., 2006. Spoken word recognition in context: evidence from Chinese ERP analyses. Brain Lang. 96 (1), 37–48. https://doi.org/10.1016/j.bandl.2005.08.007.

Luce, P.A., Pisoni, D.B., 1998. Recognizing spoken words: The neighborhood activation model. Ear Hear. 19 (1), 1–36.

MacGinitie, W.H., MacGinitie, R.K., Maria, K., Dreyer, L.G., Hughes, K.E., 2006. Gates-MacGinitie Reading Tests (GMRT). Riverside, Rolling Meadows, IL.

Malins, J.G., Joanisse, M.F., 2012. Setting the tone: an ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese. Neuropsychologia 50 (8), 2032–2043. https://doi.org/10.1016/j.neuropsychologia.2012.05.002.

Malins, J.G., Desroches, A.S., Robertson, E.K., Newman, R.L., Archibald, L.M.D., Joanisse, M.F., 2013. ERPs reveal the temporal dynamics of auditory word recognition in specific language impairment. Developmental Cognitive Neuroscience 5, 134–148. https://doi.org/10.1016/j.dcn.2013.02.005.

Malins, J.G., Gao, D., Tao, R., Booth, J.R., Shu, H., Joanisse, M.F., Desroches, A.S., 2014. Developmental differences in the influence of phonological similarity on spoken word processing in Mandarin Chinese. Brain Lang. 138, 38–50. https://doi.org/10.1016/j.bandl.2014.09.002.

Marian, V., Spivey, M., 2003a. Bilingual and monolingual processing of competing lexical items. Appl. Psycholinguist. 24, 173–193. https://doi.org/10.1017/S0142716403000092.

Marian, V., Spivey, M., 2003b. Competing activation in bilingual language processing: within- and between-language competition. Biling. Lang. Cognit. 6 (2), 97–115. https://doi.org/10.1017/s1366728903001068.

Marian, V., Blumenfeld, H., Kaushanskaya, M., 2007. The language history and experience questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. J. Speech Lang. Hear. Res. 50, 940–967. https://doi.org/10.1044/1092-4388(2007/067.

Marian, V., Spivey, M., Hirsch, J., 2003. Shared and separate systems in bilingual language processing: converging evidence from eyetracking and brain imaging. Brain Lang. 86 (1), 70–82. https://doi.org/10.1016/s0093-934x(02)00535-7.

Marinova-Todd, S.H., Zhao, J., Bernhardt, M., 2010. Phonological awareness skills in the two languages of Mandarin-English bilingual children. Clin. Linguist. Phon. 24 (4–5), 387–400. https://doi.org/10.3109/02699200903532508.

McBride-Chang, C., Tong, X., Shu, H., Wong, A.M.Y., Leung, K.W., Tardif, T., 2008. Syllable, phoneme, and tone: Psycholinguistic units in early Chinese and English word recognition. Sci. Stud. Read. 12 (2), 171–194.

McClelland, J.L., Elman, J.L., 1986. The TRACE model of speech perception. Cognit. Psychol. 18, 1–86.

McMurray, B., Samelson, V.M., Lee, S.H., Tomblin, J.B., 2010. Individual differences in online spoken word recognition: implications for SLI. Cognit. Psychol. 60 (1), 1–39. https://doi.org/10.1016/j.cogpsych.2009.06.003.

Zhao, X., Li, P., 2010. Bilingual lexical interactions in an unsupervised neural network model. Int. J. Bilingual Educ. Bilingual. 13 (5), 505–524.

Modern Chinese Frequency Dictionary, 1986. Beijing Language Institute Publisher, Beijing (in Chinese).

Naglieri, J.A., 1985. Matrix Analogies Test—Expanded Form. Psychological Corporation, San Antonio, TX.

Newman, R.L., Connolly, J.F., Service, E., McIvor, K., 2003. Influence of phonological expectations during a phoneme deletion task: evidence from event-related brain potentials. Psychophysiology 40 (4), 640–647. https://doi.org/10.1111/1469-8986.00065.

Newman, R.L., Forbes, K., Connolly, J.F., 2012. Event-related potentials and magnetic fields associated with spoken word recognition. In: Spivey, M.J., McRae, K., Joanisse, M.F. (Eds.), The Cambridge Handbook of Psycholinguistics. Cambridge University Press, New York, pp. 127–156.

Perani, D., Abutalebi, J., 2005. The neural basis of first and second language processing. Curr. Opin. Neurobiol. 15 (2), 202–206. https://doi.org/10.1016/j.conb.2005.03.007.

Powell, M.J.D., 2009. The BOBYQA Algorithm for Bound Constrained Optimization without Derivatives. Cambridge NA Report NA2009/06. University of Cambridge, Cambridge, pp. 26–46.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rueckl, J.G., Paz-Alonso, P.M., Molfese, P.J., Kuo, W.J., Bick, A., Frost, S.J., et al., 2015. Universal brain signature of proficient reading: evidence from four contrasting languages. Proc. Natl. Acad. Sci. Unit. States Am. 112 (50), 15510–15515.

Sabourin, L., Stowe, L.A., 2008. Second language processing: when are L1 and L2 processed similarly? Sec. Lang. Res. 24 (3), 397–430.

Shook, A., Marian, V., 2013. The bilingual language interaction network for comprehension of speech. Biling. Lang. Cognit. 16 (2), 304–324.

Shu, H., Chen, X., Anderson, R.C., Wu, N., Xuan, Y., 2003. Properties of school Chinese: implications for learning to read. Child Dev. 74 (1), 27–47. https://doi.org/10.1111/1467-8624.00519.

Shuai, L., Malins, J.G., 2017. Encoding lexical tones in jTRACE: a Simulation of monosyllabic spoken word recognition in Mandarin Chinese. Behav. Res. Methods 49 (1), 230–241.

Thorndike, R.L., Hagen, E.P., Sattler, J.M., 1986. Stanford-binet Intelligence Scale, fourth ed. Riverside Publishing, Chicago.

Van Heuven, W.J., Dijkstra, T., 2010. Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. Brain Res. Rev. 64 (1), 104–122.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., Parks, M., 1999. Time course of word identification and semantic integration in spoken language. J. Exp. Psychol. Learn. Mem. Cognit. 25 (2), 394–417.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., 1999. Comprehensive Test of Phonological Processing (Austin, TX: Pro-Ed. ).

Wang, X., Hui, B., Chen, S., 2020. Language selective or non-selective in bilingual lexical access? It depends on lexical tones! PLoS One 15 (3), e0230412.

Wang, X., Wang, J., Malins, J.G., 2017. Do you hear 'feather' when listening to 'rain'? Lexical tone activation during unconscious translation: evidence from Mandarin-English bilinguals. Cognition 169, 15–24. https://doi.org/10.1016/j.cognition.2017.07.013.

Wiener, S., 2019. Second language learners develop non-native lexical processing biases. Biling. Lang. Cognit. 1–12 https://doi.org/10.1017/S1366728918001165.

Wong, A.W.K., Wang, J., Wong, S.S., Chen, H.C., 2018. Syllable retrieval precedes sub-syllabic encoding in Cantonese spoken word production. PLoS One 13 (11), e0207617. https://doi.org/10.1371/journal.pone.0207617.

Woodcock, R.W., McGrew, K.S., Mather, N., 2001. Woodcock-Johnson III Tests of Achievement. Riverside Publishing, Itasca, IL.

Wu, Y.J., Thierry, G., 2010. Investigating bilingual processing: the neglected role of language processing contexts. Front. Psychol. 1, 178.

Xue, J., Yang, J., Zhang, J., Qi, Z., Bai, C., Qiu, Y., 2013. An ERP study on Chinese natives' second language syntactic grammaticalization. Neurosci. Lett. 534, 258–263.

Zhao, J., Guo, J., Zhou, F., Shu, H., 2011. Time course of Chinese monosyllabic spoken word recognition: evidence from ERP analyses. Neuropsychologia 49 (7), 1761–1770. https://doi.org/10.1016/j.neuropsychologia.2011.02.054.