The roles of lexical tone and rime during Mandarin sentence comprehension

Zou, Yun; Lui, Ming ; Tsang, Yiu Kei

Published in:
Neuropsychologia

DOI:
10.1016/j.neuropsychologia.2020.107578

Published: 01/10/2020

Document Version:
Peer reviewed version

RUNNING HEAD: CHINESE LEXICAL TONE

The roles of lexical tone and rime during Mandarin sentence comprehension: An event-related potential study

Yun ZOU, Ming LUI, Yiu-Kei TSANG

Department of Education Studies, Hong Kong Baptist University

Centre for Learning Sciences, Hong Kong Baptist University

Corresponding author:

Yiu-Kei TSANG

Department of Education Studies

Hong Kong Baptist University

Kowloon Tong, Kowloon

Hong Kong

Phone: +852 - 34115325

Email: yktsang@hkbu.edu.hk

Abstract

This study used event-related potential (ERP) recording to examine the role of lexical tone and rime in Mandarin Chinese spoken sentence comprehension. A violation paradigm was adopted, such that selected target syllables in the sentences were replaced with tone-violated, rime-violated, or double-violated syllables. Participants judged whether each sentence was congruent. The behavioral results confirmed previous findings: Tone violation was more difficult to detect than rime violation. The ERP results showed that rime and double violations, but not tone violation, elicited a larger N400 than the original condition. Similarly, tone and rime violations elicited a larger P600 than the original condition, and the effect started and ended 50 ms earlier in the tone-violation type. Interestingly, the double-violation type differed significantly from the original type only in the posterior electrodes, suggesting a weaker P600 effect than the tone- and rime-violation types. The differences in ERP effects between rime and tone processing indicate that rime played a more important role in semantic access, while tone played a more important role in error recovery. A model of Chinese speech perception was proposed to accommodate the different roles of lexical tone and rime at different processing stages during sentence comprehension.

Keywords: Lexical tone, rime, Mandarin Chinese, event-related potentials, speech perception

Introduction

Languages in the world can be divided into tonal and nontonal languages based on whether or not lexical tone constrains lexical access (Bao et al., 2013; Jaisin, Suphanchaimat, Figueroa, & Warren, 2016; Järvikivi, Vainio, & Aalto, 2010). In nontonal languages such as English and Dutch, variation in tone can reveal the emotional status of speakers, but it plays little role in constraining spoken word identity (Nygaard & Lunders, 2002). Instead, word identity is mainly determined by segmental information (Cutler & Chen, 1997). In contrast, lexical tone is vital for spoken word identification in tonal languages, such as Chinese and Thai, in which it combines with segmental information during spoken word recognition (Sereno & Lee, 2015). For example, there are four lexical tones in Mandarin Chinese—namely, the 1.) flat tone, 2.) rising tone, 3.) dipping tone, and 4.) falling tone—such that the same segment /ba/ means "to pull" in the rising tone (/ba2/) and "father" in the falling tone (/ba4/). Lexical tone also plays a similar role in Cantonese (a Chinese dialect widely spoken in Guangdong and Hong Kong). There are six tones in Cantonese, including 1.) high-level tone (/fu1/ "skin"), 2.) mid-rising tone (/fu2/ "tiger"), 3.) mid-level tone (/fu3/ "rich"), 4.) low-falling tone (/fu4/ "to support"), 5.) low-rising tone (/fu5/ "woman"), and 6.) low-level tone (/fu6/ "father"). It can be noted that in Mandarin, the four tones differ in both pitch height and pitch contour, while in Cantonese, some tones share the same contour and are differentiated only by pitch height (i.e., the three level tones and two rising tones; Tsang, Jia, Huang, & Chen, 2011).

Given the role of lexical tone in spoken Chinese, it is important to examine tonal processing in order to develop a complete model of Chinese spoken word recognition. More generally, over 40% of the world's languages are tonal languages (Jia, Tsang, Huang, & Chen, 2015). The investigation of how lexical tone plays a role in spoken word recognition is crucial to understand the universal and language-specific aspects of spoken word recognition. Acoustically, lexical tone is the fundamental frequency (f0) of speech sound, and the variation in f0 is mainly carried by the rime in Chinese (Li et al., 2014; Shuai & Gong, 2014). In other words, the availability of the acoustic signals for lexical tone and rime is tightly coupled in time. This naturally leads to the question of whether lexical tone and rime are processed simultaneously during Chinese spoken word recognition. This research question has been investigated in previous studies, but the results remain inconclusive. Some studies found that lexical tone was activated more slowly, playing a less important role than rime (Cutler & Chen, 1997; Hu, Gao, Ma, & Yao, 2012; Sereno & Lee, 2015), while others indicated that tone and rime were activated concurrently and played comparable roles during Chinese spoken word recognition (Choi, Tong, Gu, Tong, & Wong, 2017; Malins & Joanisse, 2010; Schirmer, Tang, Penney, Gunter, & Chen, 2005).

Early research on the topic relied heavily on behavioral measures such as reaction times and error rates. For example, in one experiment, Ye and Connine (1999) presented participants with a series of monosyllabic Mandarin Chinese words. At the offset of each stimulus, participants were required to judge whether it contained a specific vowel-plus-tone combination (e.g., /a/ in tone 2). The results showed that participants could make correct rejection more quickly when the syllables were mismatched in vowels (e.g., /bi2/) rather than in tones (e.g., /ba4/). Ye and Connine interpreted the results as reflecting that lexical tones were activated more slowly and constrained word identity less efficiently than vowels during Chinese spoken word recognition. Such tone disadvantage was also observed by Cutler and Chen (1997), using both the lexical decision task and the same-different judgment task in Cantonese Chinese. Accordingly, Cutler and Chen proposed that variations in f0 required more time to process than phonological segments.

On the other hand, the tone disadvantage reported above might be reduced or eliminated when a prior context is available to exert top-down modulation during spoken word recognition. For example, in another experiment, Ye and Connine (1999) showed that Mandarin tone mismatch could be detected as quickly as vowel mismatch when the target syllable was embedded in a highly constraining idiom context. Similar results were obtained by Liu and Samuel (2007) using both Mandarin sentence and idiom contexts.

The processing of tonal and segmental information during Chinese spoken word recognition has also been examined in event-related potential (ERP) studies. In contrast to behavioral indices, which reflect the sum of all processes that lead to the behavioral responses, the ERP technique has high temporal resolution that allows continuous monitoring of cognitive processes before an explicit response is made (Huang, Yang, Zhang, & Guo, 2014; Schirmer et al., 2005; Zhao, Guo, Zhou, & Shu, 2011). In addition, previous studies have established several ERP components of different latencies and polarities that are associated with different language-related processes.

One famous language-related ERP component is the N400, which is a negative component that peaks approximately 400 ms after the target stimulus onset. It has long been recognized as an index of semantic processing, such that more difficulties in semantic processing lead to larger N400 amplitudes (Barber, Otten, Kousta, & Vigliocco, 2013; Koelsch et al., 2004). Another widely studied component, the P600, is a positive component that peaks approximately 600 ms. It is related to syntactic processing and semantic reanalysis (Tanner, Grey, & van Hell, 2017; Hu et al., 2012). It is noteworthy that the N400 and P600 tend to emerge earlier in the auditory modality than the visual modality (Kutas & Federmeier, 2011). There are also slight modality differences in their topographic distributions. In particular, the auditory N400 has a more frontal and left-hemispheric distribution, while the auditory P600 is more lateralized to the right hemisphere (Osterhout & Holcomb, 1993). By measuring the latencies and amplitudes of the auditory N400 and P600, it is possible to clarify the time course of lexical tone and segment activation during Mandarin sentence comprehension.

Using the ERP technique, Hu et al. (2012) investigated tone and rime processing in Chinese idioms using the violation paradigm. In their study, participants listened to four-syllable idioms (e.g., /yi3-luan3-ji1-shi2/ “以卵击石”) and were asked to judge whether the last syllables formed correct idioms. The last syllables of the idioms could be the original syllables (e.g., /shi2/) that formed the correct idioms, or they could be the tone-violated (e.g., /shi1/), rime-violated (e.g., /shu2/), or double-violated syllables (e.g., /shu1/) that formed incorrect idioms. The behavioral results showed no differences between the tone- and rime-violation types, which is consistent with Ye and Connine (1999) and Liu and Samuel (2007). However, significant differences were observed in the ERP results. First, only the rime-violated syllables elicited a larger N100 when compared to the original syllables, which formed correct idioms. Second, while both the rime- and tone-violated syllables elicited larger N400s than the original syllables, the amplitude was significantly larger in the rime- than the tone-violation type. Third, the tone-violation type produced more robust P600s than the rime-violation type.

According to Hu et al. (2012), the N100 component in their idiom experiment reflected a lexical selection process in the early stage of spoken word recognition. Its amplitude was sensitive to whether a cohort of lexical candidates activated by the initial sound of the input contained the lexical features that matched with the expectations derived from the strong idiom context. A larger N100 amplitude would be elicited when the input did not match the expectation during this initial assessment. Given that the N100 was found only in the rime- but not the tone-violated type, the authors argued that rime was available earlier than lexical tone to match with the expectation derived from the idiom context. The larger N400 in the rime-violated type, as well as the more robust P600 in the tone-violated type suggested that while rime was more important in semantic processing, tone was more important in reanalysis after a violation had been detected. Overall, the results supported a functional dissociation between segment (rime) and lexical tone processing.

Similar results were obtained in other ERP experiments of Mandarin processing. For example, in Huang, Liu, Yang, Zhao, and Zhou (2018), participants performed a delayed syllable monitoring task after hearing idioms. Again, the effect of rime violation again preceded that of tone violation. In a picture-matching experiment, Malins and Joanisse (2012) asked participants to judge whether the spoken words they heard matched the visual objects presented. The visual objects could either match the target (e.g., /hua1/ "flower") or be mismatched in tone (e.g., /hua4/ "painting") or rime (e.g., /hui1/ "gray"). Their results showed that the tone-mismatch type produced less robust N400 effects than the rime-mismatch type, which suggested that the tone disadvantage was not restricted to idiom processing.

In contrast, these results were inconsistent with a study that employed highly constraining Cantonese sentence contexts in a violation paradigm (Schirmer et al., 2005). Participants listened to a series of Cantonese sentences that contained the original correct syllables (e.g., /beng6/; "illness"), tone-violated syllables (e.g., /beng2/; "biscuit"), rime-violated syllables (e.g., /bou6/; "step"), or he complete-violated syllables (e.g., /gwai3/; "season"). Only the sentences that contained the original words were semantically congruous, while the other types were semantically incongruous. The participants were asked to judge whether each sentence was congruous or not after its offset by pressing predefined buttons on a keyboard. The results showed that the tone- and rime-violated syllables in the incongruous sentences led to similar N400 and P600 effects. The similar ERP effects indicated that tone and rime were activated simultaneously and played comparable roles during Cantonese sentence comprehension. Given the inconsistent results of the previous ERP studies, whether lexical tone is processed similarly and with equal weight to segments in spoken word identification is worthy of further investigation.

This ERP study aimed to investigate tone and rime processing in Mandarin sentence comprehension using the violation paradigm. Because spoken words in daily speech are normally embedded in sentences instead of idioms, the results of this study should be more ecologically valid than those of idiom studies (Hu et al., 2012; Huang et al., 2018; Ye & Connine, 1999). Specifically, idioms are fixed, such that proficient language users could easily anticipate the remaining components given the first one or two syllables. In contrast, sentences have a certain degree of flexibility that the same meanings can be conveyed by different words. Therefore, the detection of tone and rime violations in idioms might involve more strategic processing, such as anticipating the upcoming syllables. Using a sentence context should reflect more natural language use, such that the online processing of lexical tone and rime might become more simultaneous, as shown in the Cantonese study by Schirmer et al. (2005). Moreover, given that Mandarin and Cantonese adopts different tonal systems, it is interesting to compare across studies to see whether the properties of the specific tonal system might affect the results. Such a study would ultimately contribute to the development of better models of spoken word recognition, which have been mostly based on research conducted with nontonal languages.

Two competing hypotheses were motivated based on previous studies. If lexical tone and rime contribute equally during Mandarin sentence comprehension, they should play similar roles in both semantic access and error reanalysis. Consequently, the tone-violated and rime-violated syllables should elicit similar N400 and P600 effects, as reported by Schirmer et al. (2005). Alternatively, if, as Hu et al. (2012) suggested, lexical tone is less important in semantic activation but plays a stronger role in reanalysis, the tone-violation type should lead to weaker N400 effects but more robust P600 effects than the rime-violation type.

Method

Participants

Forty-one right-handed native speakers of Mandarin Chinese (13 males, mean age: 25 years) were recruited for the present study. Data collection was conducted at Hong Kong Baptist University. All participants were born in mainland China and came to Hong Kong for undergraduate or postgraduate studies. At the time of testing, they had stayed in Hong Kong for less than two years. They reported high proficiency in Mandarin, and one participant also reported proficiency in Cantonese. They were paid 200 HK dollars for participation. All participants had normal or corrected-to-normal vision and intact hearing ability. None of them had any known psychological or neurological disorders. Informed consent was obtained before the experiment. The experiment was approved by the Research Ethics Committee at Hong Kong Baptist University.

Stimuli

Based on the results of pilot testing and material matching (see below), 52 sentence frames were selected, such as "艺术团演员精彩的表演让台下的____赞不

绝口" (The fantastic performance of the art troupe wins the applause of the ____ below the stage; see Table 1). Each sentence frame was paired with four types of target, including the original word (i.e., "观众", /guan1-zhong4/, audience), and three pseudowords, respectively created by violating tone (i.e., "观肿", /guan1-zhong3/), rime (i.e., "观赚", /guan1-zhuan4/) or tone-plus-rime (i.e., "观转", /guan1-zhuan3/) in the second syllable of the original word. The second syllables, instead of the first syllables, were modified to ensure that the prior context up to the first syllables would be reasonably constraining. Only the sentence frame that contained the original syllable was a semantically congruous sentence. The other three violation types produced semantically incongruous sentences. In this way, 52 congruous sentences and 156 incongruous sentences were created. The lengths of the sentences varied from 17 characters to 23 characters, and the position of the target word was around the middle of each sentence (at least four characters away from the end of each sentence) to avoid the contamination of ERP signals by the key-press responses.

[Insert Table 1 about here]

A cloze probability test was conducted to examine the strength of sentential constraints. The sentence frames (without the target syllables) were presented to 31 native Mandarin speakers who did not participate in the main experiment. They were asked to write down the first syllables that came to their mind to complete the sentences. The cloze probability of each sentence was calculated by the proportion of participants who reported the syllables that formed the original words. The average cloze probability of the original syllables was 85% (SD 0.22). The cloze probabilities of all violation types were zero.

As shown in Table 1, the target syllables were matched on homophone density (HD) and character frequency (CF) as provided by Liu, Shu, and Li (2007). The phonological frequency (PF) was also matched based on a more recent database by Sun, Hendrix, Ma, and Baayen (2018). Repeated measures analysis of variance (ANOVA) indicated that these properties were matched (ps > .10) across the four violation types. The average word frequency of the original target was 64.9 per million, while that of the pseudowords created by violating tones, rimes, or both was 0 (Sun et al., 2018).

To balance the number of congruous and incongruous sentences, another 26 congruous sentences were added into the materials as fillers. The filler sentences had similar lengths as the critical sentences (from 17 characters to 23 characters). They were not included in the analyses.

The audio recording of all sentences (52 congruous sentences, 156 incongruous sentences, and 26 filler sentences) was produced by a female native Mandarin speaker with phonetic training in Mandarin. Each sentence was produced three times, and the best token was chosen. The target syllable of each sentence was cut off and then spliced onto another token of the congruous sentence. This operation ensured that the four violation types (original vs. tone violation vs. rime violation vs. double violation) had identical sentence frames and differed only in the target syllables. Therefore, any differences across the four violation types could not be attributed to the acoustic differences of the sentence frames. Ten participants were asked to report whether any sentences they heard sounded unnatural. They reported no unnatural sentences.

This study adopted a counterbalancing design, such that each participant only heard each sentence frame once during the experiment. The items were split into four lists. Each list contained 52 experimental items, with 13 in each violation type. The pairing between sentence frame and violation type was rotated across lists such that all possible combinations were exhausted. The same set of fillers were added to each list, such that there were 39 congruous and 39 incongruous sentences in each list.

Procedure

Participants were seated in a sound-proof laboratory facing a computer screen and two speakers symmetrically placed on the left and right. In the experiment, the two speakers presented the auditory sentences bilaterally. A cross appeared at the center of the computer screen 500 ms preceding the auditory presentation of each sentence. The cross stayed on the screen until the end of the sentence. Participants were requested to fixate their eyes on the cross throughout the entire trial. Three-hundred milliseconds after the offset of the auditory sentence, a question mark "?" appeared at the center of the screen as a probe to respond. The participants were required to decide whether the sentences were comprehensible by pressing the designated keys on the keyboard. Half of the participants pressed "J" for comprehensible sentences, and "F" for incomprehensible sentences. The button assignments were counterbalanced. After responding, the next trial started automatically after a 2500-ms blank interval. Participants were instructed to blink their eyes in this interval and to withhold their eyeblinks as much as possible in other time during the trial. Overall, each trial lasted for approximately 12 seconds. Before the main experiment, there were 12 practice trials to familiarize participants with the procedure. Feedback was given to the participants after each practice trial to indicate the correctness of their responses. The whole experiment (including EEG setup and practice session) lasted for approximately 1.5 hours.

EEG Recordings

The electroencephalogram (EEG) was recorded through an eegosports system and a 64-channel cap (ANT Neuro©) placed according to the international extended 10-20 system. The Cpz electrode was used as the online reference, and the signal was rereferenced to an electrode placed on the nose tip during offline analysis. The ground electrode was inlaid in the cap between the Fpz and Fz electrodes. Horizontal EOG was monitored by electrodes placed on the outer canthus of each eye, and vertical EOG was monitored by electrodes placed above and below the left eye. The impedance of all electrodes was kept below 20 kΩ throughout the experiment. The electrical signal was recorded at a 500-Hz sampling rate.

Data Analysis

The obtained EEG data was first offline bandpass (0.05-30 Hz) filtered. Then it was segmented into epochs of -200 ms to 1000 ms relative to the onset of the target syllable embedded in each sentence. The 200-ms period before the onset of the target syllable was taken as the baseline. After baseline correction, any epochs with amplitudes over ±80 μv and error responses were excluded from the analysis. For the 41 participants, the average numbers and percentages of trials maintained in the original, the tone-violation, the rime-violation, and the double-violation types were 11.3 (87%), 9.61 (74%), 10.8 (83%), and 11.0 (84%) respectively.

The data obtained in this study were analyzed with JASP (version 0.11.1), a free open platform for statistical analysis. For the behavioral data, only the error rates were analyzed because reaction time was uninformative in the delayed-response task (i.e., participants delayed their response until the question mark "?" appeared on the screen). Repeated measures ANOVA was conducted with Violation type (original vs. tone-violation vs. rime-violation vs. double-violation) as the within-subject independent variable. For the ERP waveforms (Figure 1), the mean signal amplitude of each violation type was calculated over each successive 50-ms time window from 0 to 1000 ms after the target syllable onset. For each time window, Repeated measures ANOVA was conducted with Violation type (original vs. tone-violation vs. rime-violation vs. double-violation), Anteriority (anterior vs. central vs. posterior) and Hemisphere (left vs. right) as within-subject independent variables. For the variables Anteriority and Hemisphere, electrodes were grouped into six regions as follows: left frontal (F1, F3, F5, FC1, FC3, FC5), right frontal (F2, F4, F6, FC2, FC4, FC6), left central (C1, C3, C5, CP1, CP3, CP5), right central (C2, C4, C6, CP2, CP4, CP6), left parietal (P1, P3, P5, PO3, PO5, PO7), and right parietal (P2, P4, P6, PO4, PO6, PO8). In the analyses, the ERP amplitudes of the electrodes in each region were averaged, producing a single amplitude value for each violation type in each time window. To avoid inflation of Type

I error due to the very numerous comparisons, the false discovery rate (FDR) correction was applied (Benjamini & Hochberg, 1995). Only two or more successive time windows with significant main effects (corrected *ps* < .05) were identified as real effects and combined into larger windows for further analysis. Within the large windows, post hoc comparisons among the four violation types were conducted with Bonferroni correction. Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) was applied whenever the assumption of sphericity was violated. Uncorrected *df* and corrected *p* values were reported.

[Insert Figure 1 about here]

Results

Behavioral Data

Repeated measures ANOVA indicated a significant effect of Violation type on error rates ($F(3, 120) = 15.5$, $p < .001$, $\eta_p^2 = .28$). Post hoc comparisons revealed that tone violation elicited more errors (Mean = 16.1%, SD = .11) than the other three violation types (original: Mean = 4.3%, SD = .062, $t(40) = 5.88$, $p < .001$, Cohen's $d = .92$; rime violation: Mean = 7.7%, SD = .088, $t(40) = 3.94$, $p = .002$, Cohen's $d = .61$; double violation: Mean = 6.2%, SD = .079, $t(40) = 4.62$, $p < .001$, Cohen's $d = .72$). The original, rime- and double-violation types did not differ among themselves (all *ps* > .1).

ERP Data

The results of the successive time-windows analyses are summarized in Table 2. Two clusters of successive time windows with significant main effects could be identified. The first cluster was a negative component from 250 ms to 400 ms, while the second cluster was a positive component from 500 ms to 700 ms. Based on the latency and polarity, the early and late clusters were recognized as N400 and P600, respectively. The scalp maps of the N400 and P600 in different violation types are shown in Figure 2.

[Insert Table 2 and Figure 2 about here]

N400. There was a significant main effect of Violation type ($F(3, 120) = 8.13, p < .001, \eta_p^2 = .17$). Post hoc analyses indicated that compared with the original type, the rime- and the double-violation types produced larger N400s ($t(40) = 4.34, p = .001$, Cohen's $d = .68$ and $t(40) = 4.38, p < .001$, Cohen's $d = .68$, respectively). The tone-violation type did not differ from the original type ($p > .1$).

In addition to the significant main effect of Violation type, there was also a significant three-way interaction among Violation type, Anteriority and Hemisphere ($F(6, 240) = 2.53, p < .05, \eta_p^2 = .06$). However, further analyses indicated that the rime- and the double-violation types elicited significant N400 effects across the six regions ($ps < .01$). The tone-violation type did not differ from the original type in all regions.

P600. There was a significant main effect of Violation type ($F(3, 120) = 5.86, p < .001, \eta_p^2 = .13$). The tone- and the rime-violation types produced larger P600s than the original type ($t(40) = 3.38$, $p = .01$, Cohen's $d = .53$ and $t(40) = 3.25$, $p < .05$, Cohen's $d = .51$, respectively). No significant difference was found between the double-violation type and the original type ($p > .1$).

In addition to the main effect of Violation type, Violation type also interacted with Anteriority ($F(6, 240) = 16.8$, $p < .001$, $\eta_p^2 = .30$). Further analyses indicated that the tone and the rime-violation types elicited larger P600s than the original type in both the Central and Posterior regions ($ps < .05$), while the double-violation type elicited a larger P600 than the original type in the posterior regions only ($p < .05$).

Successive time windows analysis for P600

Although the analysis within the 500-700 ms time window showed no differences between the tone- and the rime-violation types, Figure 1 clearly showed that the early portion of the P600 appeared to be stronger in the tone-violation type. Such inspection was supported by the successive time window analysis used to identify the N400 and P600. Therefore, the results of the successive time window analysis between 500 ms to 700 ms were described below to provide more details about the differences between the tone- and rime-violation types (see Table 3 for a summary). A similar analysis procedure was also adopted by Hu et al. (2012).

[Insert Table 3 about here]

500-550 ms. There was a significant main effect of Violation type ($F(3, 120) = 5.38$, $p < .01$, $\eta_p^2 = .12$). Only the tone-violation type elicited significantly larger positivity compared with the original type ($t(40) = 3.46$, $p < .01$, Cohen's $d = .54$). The rime- and double-violation types did not differ from the original type ($ps > .1$).

550-600 ms. Again, there was a significant main effect of Violation type ($F(3, 120) = 6.16$, $p < .01$, $\eta_p^2 = .13$). The tone- and rime-violation types both elicited larger positivities compared with the original type ($t(40) = 3.74$, $p < .01$, Cohen's $d = .58$ and $t(40) = 2.89$, $p < .05$, Cohen's $d = .45$, respectively). The double-violation type did not differ from the original type ($p > .1$).

600-650 ms. Similar to the previous windows, there was a significant main effect of Violation type ($F(3, 120) = 4.73$, $p < .01$, $\eta_p^2 = .11$). The tone- and the rime-violation types both elicited larger positivities compared with the original type ($t(40) = 2.87$, $p < .05$, Cohen's $d = .45$ and $t(40) = 3.35$, $p < .05$, Cohen's $d = .52$, respectively). The double-violation type did not differ significantly from the original type ($p > .1$).

650-700 ms. There was a significant main effect of Violation type ($F(3, 120) = 3.67$, $p < .05$, $\eta_p^2 = .084$). In contrast to the previous windows, only the rime-violation type elicited a larger positive component than the original type ($t(40) = 3.16$, $p < .05$, Cohen's $d = .49$). The tone- and double-violation types did not differ significantly from the original type ($ps > .1$).

To summarize, the successive time window analyses indicated that the tone-violation type produced a significant effect in the P600 time window that started and ended 50 ms earlier than the rime-violation type. In contrast, the double-violation type did not produce reliable effects.

Discussion

This study examined the processing of lexical tone and rime in Mandarin Chinese sentence comprehension. A violation paradigm was adopted with the target syllables embedded in the middle of the sentences. Participants were asked to judge the comprehensibility of the auditory sentences they heard (i.e., congruity judgment). It was hypothesized that, if tone and rime played comparable roles during Mandarin sentence comprehension, as in the case of Cantonese (Schirmer et al., 2005), tone- and rime-violations types should elicit similar N400 and P600. There should also be no differences in behavioral error rates between the two violation types. Alternatively, as Hu et al. (2012) suggested, tone might be less important in semantic retrieval but more important in recovery when errors are detected. In this case, the tone-violation type should elicit weaker N400 but stronger P600 effects compared with the rime-violation type. The tone-violation type might also be more readily mistaken as the correct original syllables because it could be recovered easily.

The behavioral results showed that participants made more errors in the tone violation type than the rime and double violation types. The ERP results showed that, in contrast to Hu et al. (2012), there were no effects on N100, likely because the sentence context used in this study produced weaker constraints than the idiom context in their experiment. On the other hand, clear effects were observed on N400 and P600. Both the rime- and double-violation types elicited larger N400s relative to the original congruous type, while no significant difference was found between the tone-violation and the original types. Moreover, both the tone- and rime-violation types elicited larger P600s than the original type, while the difference between the double violation and the original condition on P600 was less stable (restricted to the posterior regions). The successive time window analyses further showed that the P600 effect emerged and ended 50 ms earlier in the tone-violation type than the rime-violation type. The presence of a P600 effect without a N400 effect for the tone-violation type indicated that participants were able to retrieve the correct meanings even when the incoming tones were wrong. However, they were still sensitive to the errors and attempted to correct at a later point. These results supported the hypothesis that lexical tone and rime have different roles to play during Mandarin spoken word recognition.

Given that N400 is an index of semantic processing, the larger N400 elicited by the rime-violation type compared to the original type indicated that rime plays a crucial role in semantic processing, and the rime-violated syllable could lead to difficulties in Mandarin sentence comprehension. In contrast, no significant N400 effects were produced by the tone violation type, indicating that tone has little weight in constraining word identity. Thus, participants were more likely to regard sentences that contained a tone-violated syllable as correct sentences, resulting in the more error responses in the tone-violation type. The similar N400 effects between the rime- and double-violation types also confirmed that tone-violation did not produce extra effects.

The different weights of tone and rime in semantic processing could be explained by different information values. The idea was proposed by Tong, Francis, and Gandour (2008), who found that the average number of Mandarin syllables pronounced in each tone is 310, which is much larger than the number of syllables produced by each rime (36.8). By merely knowing the tone of a syllable, listeners could have an approximately 1/310 (0.32%) chance of guessing the syllable correctly. However, the chance of guessing the syllable correctly is much higher when the available information is the rime (1/36.8 or 2.7%). Therefore, rime is a more effective cue than tone in constraining Mandarin syllable identity.

Evidence supporting the information value view can be found in previous studies using different tasks. For example, in Tong et al. (2008), Mandarin Chinese speakers heard pairs of syllables that varied in two dimensions: tone and rime. Their task was to classify the syllables according to one dimension (e.g., tone) while ignoring changes in the other dimension (e.g., rime). If one dimension has higher information value than the other in syllable processing, it should capture more attentional resources and more strongly interfere with the classification of the other dimension. The results showed that the variation in rime interfered with tone classification more than the reverse, supporting the notion that the lower information value of tone than rime has an implication in Mandarin speech processing. The lesser weight of tone in constraining word identity was also found in a study that employed the Stroop task in Mandarin Chinese (Li, Lin, Wang, & Jiang, 2013).

Earlier onset of P600 effects could be interpreted as a faster repairing of mismatched information (Gouvea, Phillips, Kazanina, & Poeppel, 2010). Therefore, the earlier P600 effects elicited by the tone-violation compared with the rime-violation type suggested that tone is recovered faster than rime in the context of Mandarin sentence comprehension. The advantage of tone in the recovery process can also be explained by the information value perspective. Specifically, as tone has lower information value than rime in Mandarin Chinese because there are just four lexical tones, listeners only need to try three alternative tones at most to recover a tone-violated syllable. In contrast, they have to search through a larger number of alternative rimes to recover a rime-violated syllable. Apparently, recovering a tone-violated syllable requires less effort than recovering a rime-violated syllable, which might lead to a faster reanalysis of tone during Mandarin sentence comprehension. Although the syllables mismatched on a single dimension (tone or rime) can be swiftly updated under the sentence context, repairing the double-violated syllable is much harder because it misses most acoustic cues (only the onset was correct) that are expected based on the prior context. Therefore, participants might judge the sentence containing a double-violated syllable as incomprehensible once they encounter difficulties in semantic processing, without updating the syllable to match the context. Consequently, no robust P600 effects could be elicited by the double violation type.

The advantage of tone in recovery might also be related to the abundant daily experience in tone normalization. Acoustically, tone is an f0 variation. Its production can be easily affected by factors other than lexical tones (Schirmer et al., 2005; Shao & Zhang, 2019). For example, the same tone is realized differently by men and women because of the anatomical differences in the vocal cords, but listeners can retrieve the correct tone heard by compensating for the gender difference in the f0 range. Similarly, listeners can also normalize the lexical tones spoken in different moods (e.g., higher f0 in a happy mood and lower f0 in a sad mood). In addition, the tone sandhi rule can influence lexical tone production, such that when reading aloud the Mandarin word /ni2hao3/ ("hello"), the first syllable /ni2/ is pronounced in tone 2 even though it is pronounced in tone 3 (/ni3/) in isolation. To ensure smooth comprehension, listeners have to process tonal information more flexibly (Schirmer et al., 2005). The abundant experience in tone normalization might facilitate the recovering of tone in the sentence context.

The advantage of tone in the recovering stage in this study was also consistent with an earlier behavior study conducted by Wiener and Turnbull (2016). In their study, Mandarin speakers were presented with a series of pseudosyllables (e.g., /su3/). They were required to reconstruct each syllable into a correct Mandarin syllable by changing either the consonant (e.g., /tu3/), the tone (e.g., /su4/), or the rime (e.g., /si3/) of that syllable. The results showed that changing the tone led to faster and more accurate responses compared with changing the rime. Moreover, when participants were free to choose how to reconstruct the pseudosyllables, they preferred to restore them through changing the tones. These findings supported that tone violation can be recovered quickly.

The differences between processing lexical tones and rimes observed in the present sentence comprehension experiment are mostly consistent with the Mandarin idiom experiments by Hu et al. (2012) and Huang et al. (2018). In other words, the functional dissociation between tones and rimes is unlikely to be solely attributable to the strategic processing of highly constraining idioms. Instead, it is a general property of processing Mandarin Chinese speech. However, the results appear to be inconsistent with the results of Schirmer et al. (2005), which studied Cantonese sentences with Hong Kong university students.

The different results may have been due to the different language-teaching methods adopted in mainland China and Hong Kong. In mainland China, Mandarin speakers are taught *pinyin*, an alphabetic system of Chinese. In *pinyin*, tone and segment are distinct from each other (Shu, Peng, & McBride-Chang, 2008; Taft & Chen, 1992; Wang, Li, & Lin, 2015). Through formal instruction in *pinyin*, Mandarin speakers might realize the independence of the tone and segment unit, and easily notice the low information of tone in spoken word identification. Conversely, Cantonese speakers in Hong Kong learn orthography-to-pronunciation mapping without learning a similar alphabetic system. As a result, they might treat lexical tone and rime as a whole unit (McBride, Bialystok, Chong, & Li, 2004; Taft & Chen, 1992). This possibility was partly supported by a previous study that used a primed naming task, which indicated that *pinyin* primes effectively promoted the independent processing of tone and segment during spoken word planning (Wang et al., 2015). Nevertheless, it is important for further study to examine whether *pinyin* learning could truly modulate the relative weight of lexical tone and segment during sentence comprehension.

Another possible explanation of the differences between Mandarin and Cantonese centered upon the differences in the information value of lexical tone. As mentioned previously, Mandarin only has four tones, while Cantonese has six. There are also more possible syllable-tone combinations in Cantonese (1700 to 2300) than Mandarin (approximately 1300). These may have encouraged Cantonese users to extract and utilize lexical tone information as soon as possible to increase the speed of spoken word recognition. This possibility can be tested by manipulating the probability of specific syllable-tone combinations. For example, in a visual-world paradigm experiment, Wiener and Ito (2014) presented participants four visual characters and a spoken syllable simultaneously in each trial. Participants were instructed to move the mouse and clicked on the character that corresponded to the spoken syllable. The fixation pattern and mouse click response times were faster when the spoken syllables were pronounced with highly probable tones (e.g., /zhou1/ "porridge") than improbable ones (e.g., /zhou2/ "axle"), indicating that more probable tones might be more constraining. Future studies can examine the effect of tonal probability in sentence comprehension with ERP recording.

Theoretical implications

Most existing models of spoken word recognition (e.g., Cohort Model, TRACE, and the Neighborhood Activation Model) were established based on nontonal languages, such as English and Dutch (Marslen-Wilson, 1987; McClelland & Elman, 1986; Luce & Pisoni, 1998). These models only consider how segmental information is used for lexical access and meaning retrieval. However, more than 40% of the world's spoken languages are tonal (Jia et al., 2015). Therefore, understanding how lexical tone contributes to spoken word recognition is crucial to develop more universal models of spoken word recognition. In an early attempt to model spoken word recognition in tonal languages, Ye and Connine (1999) modified the TRACE model based on the results of their behavioral experiments. In the modified model, lexical tones are represented as "tonemes", an independent layer of representation from segments (i.e., phonemes). The inputted acoustic signals activate the corresponding tonemes and phonemes, which in turn combine to activate the lexical representations. Moreover, tonemes and phonemes receive feedback excitation from the lexical representations, and this feedback is assumed to be stronger for tonemes. As a result, when isolated Chinese spoken words are recognized, the activation of tonemes would lag behind phonemes because there is little top-down feedback. In contrast, the feedback becomes stronger when a specific word is expected based on the previous context, which strengthens the activation of tonemes. This explains why phonemes and lexical tones are equally available in the idiom context.

The separate representations of lexical tone and segment in Mandarin was also endorsed by the formal computer model TRACE-T (Shuai & Malins, 2017). In particular, Mandarin consonants were coded in three dimensions namely, voicing, place of articulation, and manner of articulation. Similarly, vowels were coded by roundedness, tongue position, and tongue height. Lexical tones were represented by 15 units coded on pitch height and pitch slope. Moreover, to emphasize the temporally unfolding nature of lexical tone, each time-normalized tone was divided into five intervals. The consonants, vowels, and tone units were then activated and combined incrementally as the spoken syllable was heard. Adopting such architecture, the model successfully simulated the results of two visual world paradigm experiments of Mandarin spoken word recognition (Malins & Joanisse, 2010; Wiener & Ito, 2014).

More recently, other researchers (Choi et al., 2017; Tong et al., 2014) proposed the TTRACE model, which assumes that vowels and tones are represented and processed as holistic units in tonal languages. In other words, there are multiple representations for a vowel, with one for each tone. Evidence for this model came mainly from perception-oriented experiments, such as the passive oddball paradigm. For example, Choi et al. (2017) presented participants with a sequence of "standard" syllable-tone combinations (e.g., /u1/). Occasionally, a "deviant" sound would be presented, and an ERP component called mismatch negativity (MMN) would be triggered, even when participants were not attentively detecting the deviants. The deviants could mismatch in tone (e.g., /u3/), vowel (e.g., /i1/), or both (e.g., /i3/). The results showed that the MMN generated for the double mismatch type was smaller than the sum of the MMNs for tone and vowel mismatches. This underadditive effect suggested that there was perceptual integration of vowels and lexical tones and was supportive to the holistic representation view of TTRACE.

The results of this study have two implications for modeling spoken word recognition in tonal languages. First, the different effects produced by rime- and tone-violations are more consistent with the assumption of separate representations of vowels and tones in the modified TRACE model (Ye & Connine, 1999). It is because the holistic view of TTRACE (Choi et al., 2017) implies that both types of violation would result in the same failure in activating the correct vowel-tone representation. Second, the different patterns between the N400 and P600 emphasize the importance of considering the processing of vowels and lexical tones in a more dynamic manner. Specifically, the modified TRACE and TTRACE models focused on describing how vowels and tones contributed to lexical access. On the other hand, vowels appear to be more important in semantic retrieval, while lexical tones contribute more to reanalysis for recovery from errors. This can be summarized as a dynamic model of Mandarin spoken word recognition, as shown in Figure 3.

[Insert Figure 3 about here]

According to this model, Mandarin spoken word processing in a sentence context occurs in two stages: a semantic processing stage and a reanalyzing stage. Segmental information (vowel/rime) plays the dominant role in the semantic processing stage, while tone has little weight in semantic processing. In the reanalyzing (recovery) stage, although both tone and segment are critical for repairing the mismatched spoken word, it is easier to reanalyze tone than segment because of its information value and the experience of tone normalization (e.g., gender and emotion) in daily life. Future research can examine the factors that may modulate the dynamics, such as the complexity of the lexical tone system (e.g., Cantonese has six tones and shows a more important role of tone in semantic access; Schirmer et al., 2005).

Conclusion

The present ERP study was conducted to examine the processing of rimes and lexical tones in Mandarin spoken word recognition in a sentence context. The results showed that rimes were more important than tones in semantic processing, as indicated by the stronger N400 effects. In contrast, lexical tones offered advantages in recovering errors. A dynamic model of Mandarin spoken word recognition was proposed to account for these findings. Future research can examine how factors like the information value of specific segment-tone combinations and pinyin teaching may modulate tonal processing. This kind of research in tonal languages will be necessary for understanding the processes that support spoken word recognition.

Acknowledgement

This work was supported by General Research Fund (No: 12606717) from the Research Grants Council of the Hong Kong Special Administrative Region, China.

References

Bao, Y., Szymaszek, A., Wang, X., Oron, A., Pöppel, E., & Szelag, E. (2013). Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. *Cognition, 129(3),* 579–585.

Barber, H. A., Otten, L. J., Kousta, S.-T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language, 125(1),* 47–53.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57(1),* 289–300.

Choi, W., Tong, X., Gu, F., Tong, X., & Wong, L. (2017). On the early neural perceptual integrality of tones and vowels. *Journal of Neurolinguistics, 41,* 11–23.

Cutler, A., & Chen, H.-C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics, 59(2),* 165–179.

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes, 25(2),* 149–188.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24,* 95–112.

Hu, J., Gao, S., Ma, W., & Yao, D. (2012). Dissociation of tone and vowel processing in Mandarin idioms: Tone and vowel processing in Mandarin. *Psychophysiology, 49(9),* 1179–1190.

Huang, X., Liu, X., Yang, J.-C., Zhao, Q., & Zhou, J. (2018). Tonal and vowel information processing in Chinese spoken word recognition: An event-related potential study. *NeuroReport, 29(5),* 356–362.

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.

Jaisin, K., Suphanchaimat, R., Figueroa Candia, M. A., & Warren, J. D. (2016). The Speech-to-Song Illusion Is Reduced in Speakers of Tonal (vs. Non-Tonal) Languages. *Frontiers in Psychology, 7,* 662.

Järvikivi, J., Vainio, M., & Aalto, D. (2010). Real-Time Correlates of Phonological Quantity Reveal Unity of Tonal and Non-Tonal Languages. *PLoS ONE, 5(9):* e12603.

Jia, S., Tsang, Y.-K., Huang, J., & Chen, H.-C. (2015). Processing Cantonese lexical tones: Evidence from oddball paradigms. *Neuroscience, 305,* 351–360.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7(3),* 302–307.

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology, 62(1),* 621–647.

Li, C., Lin, C. Y., Wang, M., & Jiang, N. (2013). The activation of segmental and tonal information in visual word recognition. *Psychonomic Bulletin & Review, 20(4),* 773–779.

Li, W., Wang, L., & Yang, Y. (2014). Chinese Tone and Vowel Processing Exhibits Distinctive Temporal Characteristics: An Electrophysiological Perspective from Classical Chinese Poem Processing. *PLoS ONE, 9(1):* e85683.

Liu, S., & Samuel, A. G. (2007). The role of Mandarin lexical tones in lexical access under different contextual conditions. *Language and Cognitive Processes, 22(4),* 566–594.

Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods, 39(2),* 192-198.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear and Hearing, 19,* 1–36.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language, 62(4),* 407–420.

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.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25(1–2),* 71–102.

McBride-Chang, C., Bialystok, E., Chong, K. K. Y., & Li, Y. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology, 89(2),* 93–111.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18(1),* 1–86.

Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition, 30(4),* 583-593.

Osterhout, Lee, and Phillip J. Holcomb. (1993). Event-Related Potentials and Syntactic Anomaly: Evidence of Anomaly Detection during the Perception of Continuous Speech. *Language and Cognitive Processes, 8(4),* 413–437.

Schirmer, A., Tang, S.-L., Penney, T. B., Gunter, T. C., & Chen, H.-C. (2005). Brain Responses to Segmentally and Tonally Induced Semantic Violations in Cantonese. *Journal of Cognitive Neuroscience, 17(1),* 1–12.

Sereno, J. A., & Lee, H. (2015). The Contribution of Segmental and Tonal Information in Mandarin Spoken Word Processing. *Language and Speech, 58(2),* 131–151.

Shao, J., & Zhang, C. (2019). Talker normalization in typical Cantonese-speaking listeners and congenital amusics: Evidence from event-related potentials. *NeuroImage: Clinical, 23,* 101814.

Shu, H., Peng, H., & McBride-Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science, 11(1),* 171–181.

Shuai, L., & Gong, T. (2014). Temporal relation between top-down and bottom-up processing in lexical tone perception. *Frontiers in Behavioral Neuroscience*, 8:97.

Shuai, L., & Malins, J. G. (2016). Encoding lexical tones in jTRACE: a simulation of monosyllabic spoken word recognition in Mandarin Chinese. *Behavior Research Methods*, 1-12.

Sun, C. C., Hendrix, P., Ma, J., & Baayen, R. H. (2018). Chinese lexical database (CLD): A large-scale lexical database for simplified Mandarin Chinese. *Behavior Research Methods, 50(6),* 2606–2629.

Taft, M., & Chen, H.-C. (1992). Judging Homophony in Chinese: The Influence of Tones. *Advances in Psychology, 90,* 151–172.

Tanner, D., Grey, S., & van Hell, J. G. (2017). Dissociating retrieval interference and reanalysis in the P600 during sentence comprehension: Retrieval and reanalysis in sentence comprehension. *Psychophysiology, 54(2),* 248–259.

Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes, 23(5),* 689–708.

Tong, X., McBride, C., & Burnham, D. (2014). Cues for Lexical Tone Perception in Children: Acoustic Correlates and Phonetic Context Effects. *Journal of Speech Language and Hearing Research, 57(5),* 1589–1605.

Tsang, Y.-K., Jia, S., Huang, J., & Chen, H.-C. (2011). ERP correlates of pre-attentive processing of Cantonese lexical tones: The effects of pitch contour and pitch height. *Neuroscience Letters, 487(3),* 268–272.

Wang, M., Li, C., & Lin, C. Y. (2015). The Contributions of Segmental and Suprasegmental Information in Reading Chinese Characters Aloud. *PLoS ONE*, *10(11):* e0142060.

Wiener, S., & Ito, K. (2014). Do syllable-specific tonal probabilities guide lexical access? Evidence from Mandarin, Shanghai and Cantonese speakers. *Language, Cognition and Neuroscience, 30(9),* 1048–1060.

Wiener, S., & Turnbull, R. (2016). Constraints of Tones, Vowels and Consonants on Lexical Selection in Mandarin Chinese. *Language and Speech, 59(1),* 59–82.

Ye, Y., & Connine, C. M. (1999). Processing Spoken Chinese: The Role of Tone Information. *Language and Cognitive Processes, 14(5–6),* 609–630.

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.

Table 1. Sample stimuli and lexical properties across violation types.

Sample sentence frame: 艺术团演员精彩的表演让台下的**观**__赞口不绝

(The fantastic performance of the art troupe wins the applause of the ____ under the stage)

Type	Original	TV	RV	DV
Syllable	众, “crowd’	肿, “swelling”	赚, “earn”	转, “transfer”
Pronunciation	/zhong4/	/zhong3/	/zhuan4/	/zhuan3/
HD	22.3 (17.7)	26.6 (32.8)	22.9 (17.6)	20.0 (20.6)
PF	2.70 (0.55)	2.78 (0.77)	2.82 (0.71)	2.54 (0.75)
CF	1.21(0.73)	1.33(0.67)	1.35(0.91)	1.32(0.73)
Duration*	229 (42.4)	247 (37.5)	245 (42.1)	249 (42.4)

Note. TV: tone violation; RV: rime violation; DV: double violation; HD: homophone density; PF: phonological frequency after logarithmic transformation; CF: character frequency after logarithmic transformation; *Duration was not matched because the original syllable was shorter than tone, rime and double violated syllables (ps < .01), which did not differ among themselves (ps > .10); Standard deviation (SD) is put in the parentheses.

Table 2. *F* Values of the Violation type × Hemisphere × Anteriority ANOVA.

Time windows (ms)	Violation type	Violation type x Hemisphere	Violation type x Anteriority	Violation type x Anteriority x Hemisphere
0-50	0.41	1.27	1.36	0.51
50-100	0.89	0.60	1.25	0.43
100-150	1.83	0.84	0.77	0.56
150-200	2.59	0.99	0.76	0.51
200-250	2.11	0.70	1.33	0.81
250-300	6.14**	0.29	0.70	2.39
300-350	6.78**	1.24	0.67	2.08
350-400	8.06**	0.56	0.92	2.00
400-450	2.76	1.16	2.21	1.10
450-500	1.70	0.86	5.74**	1.11
500-550	5.38**	1.57	6.94**	1.27
550-600	6.16**	0.41	14.3**	1.52
600-650	4.73*	2.89	17.0**	1.68
650-700	3.67*	2.54	21.9**	2.52
700-750	2.61	4.22	18.5**	0.61
750-800	2.23	1.73	16.0**	1.08
800-850	2.27	2.91	12.7**	1.08
850-900	2.38	2.10	10.0**	1.40
900-950	1.11	2.15	9.75**	0.90
950-1000	0.80	0.85	7.15**	0.82

**p* < .05; ***p*< .01 (The *p*-values are FDR corrected).

Table 3. *A summary of the details of P600 effects*

Time window (ms)	TV vs. Original	RV vs. Original	DV vs. Original
500~550	TV > Original	NS	NS
550~600	TV > Original	RV > Original	NS
600~650	TV > Original	RV > Original	NS
650~700	NS	RV > Original	NS

Note. TV: tone violation; RV: rime violation; DV: double violation.

Figure captions

Figure 1. Average ERP waveforms in different violation types.

Figure 2. Scalp maps of N400 and P600 effects in the tone violation (TV-O), the rime violation (RV-O), and the double violation type (DV-O).

Figure 3. A dynamic model of Mandarin spoken word processing. Thicker lines represent stronger connections.

Figure 1. Average ERP waveforms in different violation types.

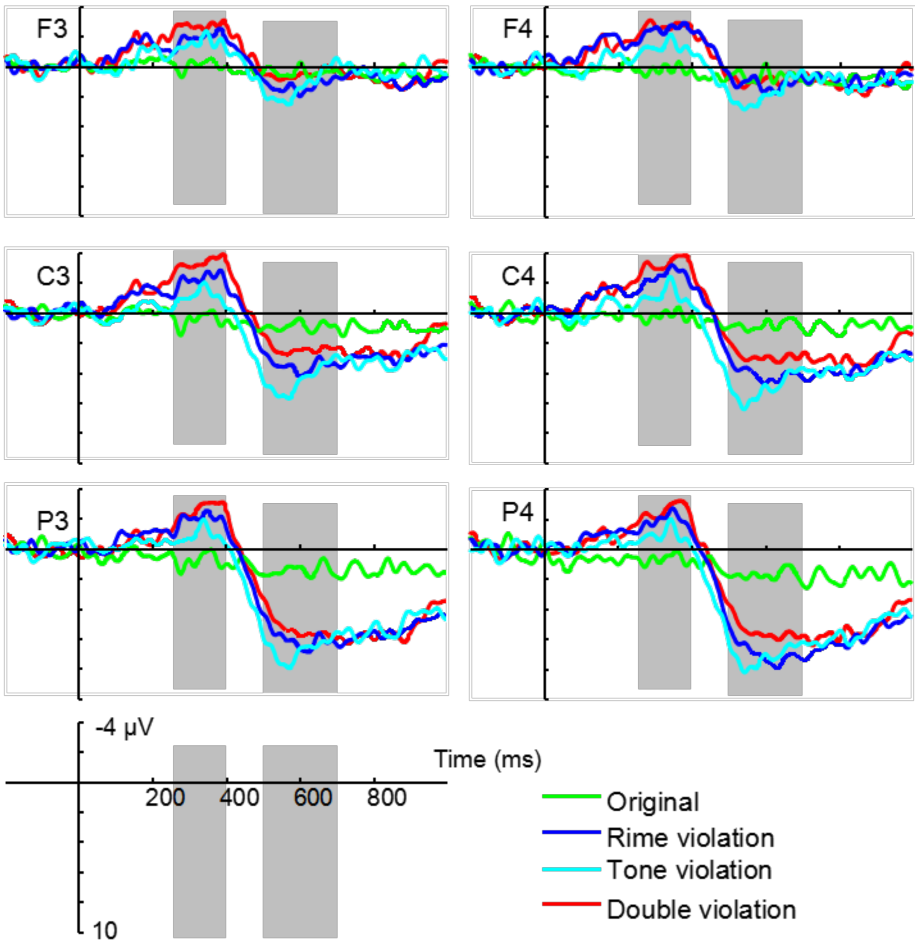

Figure 2. Scalp maps of N400 and P600 effects in the tone violation (TV-O), the rime violation (RV-O), and the double violation type (DV-O).

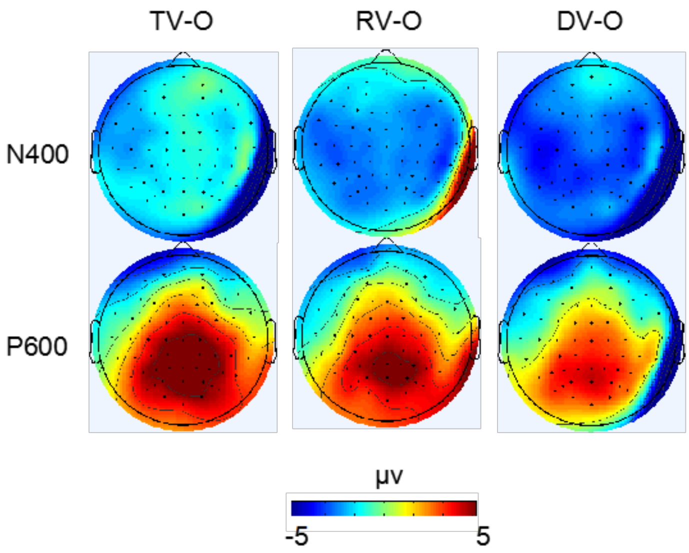

Figure 3. A dynamic model of Mandarin spoken word processing. Thicker lines represent stronger connections.