12-3-2019

Integrated vs independent processing of lexical tone and rime in mandarin sentence comprehension : an event-related potential study

Yun Zou

HONG KONG BAPTIST UNIVERSITY

Master of Philosophy

THESIS ACCEPTANCE

DATE: December 3, 2019

STUDENT'S NAME: ZOU Yun

THESIS TITLE: Integrated vs Independent Processing of Lexical Tone and Rime in Mandarin Sentence Comprehension: An Event-related Potential Study

This is to certify that the above student's thesis has been examined by the following panel members and has received full approval for acceptance in partial fulfilment of the requirements for the degree of Master of Philosophy.

Chairman:	Prof Leung Allen Y L Professor, Department of Education Studies, HKBU (Designated by the Dean of Faculty of Social Sciences)
Internal Members:	Prof Tam Vicky C W Professor, Department of Education Studies, HKBU (Designated by the Head of Department of Education Studies)
	Dr Tsang Yiu Kei Associate Professor, Department of Education Studies, HKBU
External Examiner:	Dr Wong Wing Kuen Associate Professor Department of Applied Social Sciences City University of Hong Kong

Issued by Graduate School, HKBU

Integrated vs Independent processing of lexical tone and rime in Mandarin sentence comprehension: An event-related potential study

ZOU Yun

A Thesis Submitted in partial fulfilment of the requirements

for the degree of

Master of Philosophy

Principal Supervisor:

Dr. TSANG Yiu Kei (Hong Kong Baptist University)

December 2019

DECLARATION

I hereby declare that this thesis represents my own work which has been done after registration for the degree of MPhil at Hong Kong Baptist University, and has not been previously included in a thesis or dissertation submitted to this or any other institution for a degree, diploma or other qualifications.

I have read the University's current research ethics guidelines, and accept responsibility for the conduct of the procedure in accordance with the University's Research Ethics Committee (REC). I have attempted to identify all the risks related to this research that may arise in conducting this research, obtained the relevant ethical and safety approval, and acknowledged my obligations and the rights of the participants.

Signature: ZOU YUN

Date: December 2019

Abstract

Although tone plays an important role in Chinese speech perception, many issues regarding tone processing remained debatable. Among these issues, whether tone and segment are integrally or independently processed is the one that motivated this study. To investigate the integration vs. independence view of tone and segment (i.e., rime in this study) processing during Mandarin sentence comprehension, the present experiment adopted the violation paradigm and event-related potential (ERP) recording. Participants heard a series of sentences and judged whether each sentence made sense or not after its offset. The sentences included congruous sentences that embodied the original word (e.g., “观众”/guan1-zhong4/; “audience”) and incongruous sentences that were created by mismatching tone (e.g., “观肿”/guan1-zhong3/), rime (e.g., “观赚”/guan1-zhuan4/) or tone-plus-rime (e.g., “观转”/guan1-zhuan3/) of the second syllable of the original words in the congruous sentences. Larger N400 (250-400 ms) and P600 (500-700 ms) were evoked by the incongruous sentences than congruous sentences. Among the incongruous sentences, the N400 elicited by double violation (i.e., tone-plus-rime violation) was larger than rime violation, which was in turn larger than tone violation. The P600 evoked by tone violation was larger than rime violation in 500-600 ms, but they were comparable in 600-700 ms. In addition, the P600 evoked by tone and rime violation were both larger than double violation. The different ERP effects among the three violation conditions supported the independence view of tone and rime processing. Based on the results, a dynamic model of spoken word processing was proposed, in which the functional dissociation of tone and segment across different stages was taken into consideration.

Acknowledgement

I want to thank my supervisor Dr. TSANG YIU KEI for his valuable comments and continuous encouragement in conducting the research. Thanks are also given to Dr. LUI Ann, Mr. LAU Gilbert, and all my thesis committee members for their assistance and advice.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Languages in the world can be divided into tonal and non-tonal languages based on whether or not tone carries lexical meanings in spoken words. Many Indo-European languages such as English and Dutch, are non-tonal, meaning that tone is not exploited to distinguish word identity. On the contrary, tone can determine word meaning and is indispensable for spoken word identification in tonal languages, such as Chinese and Thai. In Mandarin Chinese, there are four tones in total, including a high level tone (tone 1); a high rising tone (tone 2); a low dipping tone (tone 3); and a high falling tone (tone 4). The same segment pronounced in different tones can mean totally different things. For instance, the segment 'ba' combining with tone 1 would mean 'eight', while it means 'pull', 'handle', 'father' when it combines with tone 2, tone 3, tone 4, respectively.

Abundant studies have justified the importance of tone in Chinese speech processing (Brown-Schmidt & Canseco-Gonzalez, 2004; Lee, 2007; Li, Wang, & Yang, 2014; S. Liu & Samuel, 2007; Malins & Joanisse, 2010; Schirmer, Tang, Penney, Gunter, & Chen, 2005; X. Tong, McBride, & Burnham, 2014). Yet, many details remained unsettled. In particular, whether tone and segment are integrally or independently processed during spoken word identification has evoked vigorous debates. To date, some studies found that tone and segment are integrally processed as a unified processing unit (Choi & Tong, 2016; Gao et al., 2012; Singh, Lee, & Goh, 2011), while others indicated that they are separately processed as independent processing units (Hu, Gao, Ma, & Yao, 2012; Li et al., 2014; Luo et al., 2006; Ye & Connine, 1999). It is worth noting that most of these studies only considered tone and segment processing during monosyllable perception, while few investigated their processing in sentence context, which is

the most realistic context for spoken word processing in daily speech. The knowledge of integration or independence between tone and segment processing would remain incomplete without uncovering their processing in sentence level. To fill this gap, the present study is conducted to explore tone and segment (rime) processing in the context of Mandarin sentence comprehension, which would be a critical step for constructing more valid theoretical models of Chinese spoken word identification. Specifically, a spoken sentence comprehension experiment is set up to examine whether lexical tone and rime are processed as a unified unit or as independent units.

1.1 Previous studies exploring the integration or independence between tone and segment processing

Acoustically, tone and segment are defined by different features, namely fundamental frequency and format frequencies, respectively. Yet, variation of fundamental frequency is carried by vowel. In this sense, tone and segment "overlap" with each other and contribute to spoken word identification together. Then, a question naturally arises regarding their relationship during Chinese spoken word recognition: "Are they processed integrally or separately?" This question is of high theoretical importance, as it is a fundamental issue in modelling Chinese spoken word identification.

There are two major, competing views that describe how lexical tones are activated and combined with segment in the mental lexicon. Although the models make different assumptions about the representations and processing of lexical tone, they are both developed based on the TRACE model (McClelland & Elman, 1986), a connectionist theory of English spoken word recognition. According to the original TRACE model, there are three layers of representation,

namely the feature layer, phoneme layer, and word layer. In addition to the feedforward connection from features to phonemes and then to words, feedback connection also exists from words to phonemes.

The first model that incorporates lexical tone representation to understand Chinese spoken word processing is the "modified TRACE model" proposed by Ye and Connine (1999). The model retains the three layers of representation in the original TRACE model. Moreover, it assumes that lexical tone is represented as a separate "toneme" layer, which is independent of the phoneme layer. The processing system extracts the fundamental frequency of the input acoustic features and activates the corresponding tonemes. When a toneme is activated, it sends feedforward connection to the word (i.e., a syllable in Chinese) layer and combines with the feedforward input from the phoneme layer to activate the corresponding word representation. Feedback connection exists between words and phonemes, as well as between words and tonemes. However, the feedback is assumed to be stronger in the latter case. As tone has its own representation which can be activated independently from segment in the model, tone and segment are considered to be processed separately from each other.

The second model is the TTRACE model developed by X. Tong, McBride and Burnham (2014). In contrast to the modified TRACE model, TTRACE assumes a phonological layer of representation that integrates phonemes and lexical tones. In other words, in the phonological layer, every vowel has multiple representations, each of a different tone. Accordingly, lexical tone does not have an independent representation. An integrated "vowel-plus-tone" representation is activated based on the input acoustic features, which combines with the consonant at the phonological layer to activate the word layer representations. Since tone and

segment (i.e., vowel) are integrally represented in the phonological layer, they are considered to be integrally processed as a unified processing unit ("vowel-plus-tone") according to the model.

To date, both models received supports from different evidence, whether tone and segment are integrally or independently processed remained inconclusive.

1.1.1 Evidence supporting the independent view of modified TRACE model

If tone and segment are integrally processed as the same processing unit, they should be activated at similar time point and be equally important to lexical activation. On the other hand, if tone and segment are independently processed, they could have different weights in constraining word identity and there could be a temporal dissociation between their activation. Therefore, by examining the relative importance of tone and segment as well as the timing of their activation, it would be possible to answer the question of whether they are integrally or separately processed during spoken word identification.

Ye and Connine (1999) found a temporal dissociation between tone and segment activation in monosyllable word perception, which was in favor of the independent view of the modified TRACE model. In their study, the participants were presented with Mandarin Chinese monosyllabic stimuli. Half of them were real words (e.g., /ba2/) and the other half were pseudowords (e.g., /ra2/) that sounded like real words but did not correspond to a meaningful real word. Participants' task was to judge whether the word they heard contain a particular tone-plus-vowel combination (e.g., /a2/). They found that the participants rejected the stimulus that mismatched in vowel (e.g., /bi2/) faster than that mismatched in tone (e.g., /ba4/) for both real words and pseudowords. The results indicated that vowel is activated faster than tone during spoken word perception. The temporal dissociation between tone and segment (vowel)

activation found in the study could be explained by the independent view of modified TRACE model that tone and segment are independently activated.

The independent view of modified TRACE model also received supports from a spoken word production study by Wong and Chen (2015), which showed that tone and segment are separately prepared during spoken word production. In their study that adopted the picture-word interference (PWI) paradigm, Cantonese-speaking participants were asked to name aloud pictures whose names are all Cantonese monosyllables with a consonant-vowel-consonant structure (e.g., 星 /sing1/, "Star") and ignore an accompanying auditory word distractor. The distractor words included: 1) a monosyllable sharing two identical segments with the target (e.g., /soeng3/); 2) a monosyllable sharing only one identical segment with the target (e.g., /hung2/); 3) a monosyllable sharing one identical segment plus tone with the target (e.g. /fung1/); 4) a monosyllable sharing two identical segments plus tone with the target (e.g., /soeng1/); 5) unrelated control monosyllable (e.g., /gok3/). The results showed that the distractor sharing two identical segments with the target facilitated the target word naming, whereas the distractor sharing only one identical segment or one segment plus tone produced no significant effects. Besides, the facilitation effects of the distractor sharing two identical segments plus tone were stronger than that sharing two identical segments. Since the distractor sharing two identical segments facilitated target word naming while the distractor sharing one segment plus tone had no significant effects, it indicated that tone is not processed like a segment unit and have less weight than segment in spoken word production. In addition, the larger facilitation effects from the distractor sharing two identical segments plus tone than the distractor sharing two segments implied that tone is also activated, although it is not as effective cue as segment during spoken

word planning. Overall, the results were consistent with the claim of modified TRACE model that tone and segment are activated separately as different processing units.

In addition to behavioral measures, another popular approach to explore the integration or independence between tone and segment processing is to examine brain activity through event-related potential (ERP). ERP recording is a tool of high temporal resolution (Schirmer et al., 2005). Compared with behavioral measures, it can provide more information about how speech processing unfolds online as well as the neuronal mechanism underlying the processing. Several ERP components characterized by polarity, latency and scalp distribution have been intensively applied in speech processing. For example, the mismatch negativity (MMN) is an ERP component which peaks at 100-250 ms from change onset and exhibits the strongest intensity in temporal and frontal areas of topographic scalp maps (Luo et al., 2006). It can reflect an automatic neuronal response to changes. It is typically obtained in an oddball paradigm by presenting an oddball event (deviant) in a stream of repeated or familiar events (standards) and then subtracting the event-related response to standards from the response to deviants (Bishop & Hardiman, 2010).

Adopting the passive oddball paradigm with Mandarin monosyllable materials, Luo et al. (2006) found that the detection of tone and segment change involved different brain regions. In their study, Mandarin Chinese speakers heard a meaningful word (e.g., /bai1/, “split”) as standards and a deviant word varying either tone (e.g., /bai2/, “white”) or initial consonant (e.g., /sai1/, “plug”) in the standards. They found that the MMN in response to the lexical tone variation was stronger in magnitude when recorded on the right side of scalp than when recorded on the left side, while the MMN in response to the initial consonant variation was stronger in magnitude when recorded on the left side of scalp than on the right. The opposite lateralization

pattern of tone and initial consonant indicated that tone and segment might be processed separately in brain, which supported the independent view of modified TRACE model that tone and segment have their own separate representations.

1.1.2 Evidence supporting the integrated view of TTRACE model

While there are abundant studies supporting the independent view of modified TRACE, the integrated view of TTRACE model also received supports from behavioral and event-related potentials (ERP) evidence.

To determine whether the two sources of information (e.g., segment and tone) are processed integrally or independently, there is an influential paradigm known as Garner's (1970, 1974) speeded classification task. In this paradigm, participants listen to a series of stimuli that vary along two dimensions and classify them according to the target dimension while trying to ignore the non-target dimension. If one dimension is processed integrally with the other (i.e., the processing of one dimension entails the processing of the other dimension), listeners would have difficulty selectively attending to only one dimension. Lee and Nusbaum (1993) adopted this paradigm to investigate the integration or independence between tone and word initial consonant processing. In their study, participants were presented with Mandarin monosyllables and were asked to classify the monosyllables (e.g., /ba/ in high tone, /ba/ in low tone, /da/ in high tone, /da/ in low tone) according to either tone (e.g., high vs. low tone) or word initial consonant (e.g., b vs. d) as the target dimension. The results showed a symmetric interference between tone (e.g., high vs. low tone) and word initial consonant (e.g., /b/ vs. /d/) processing. Specifically, when the target dimension was tone (i.e., determining whether the target syllable they heard was of a high or low tone), the variation in word initial consonant (e.g., /ba/ - /da/) could inhibit tone

classification (e.g., determine the tone of the word is high or low) significantly and vice versa, indicating that tone and segment (consonant) are integrally processed with each other. The integration between tone and segment processing found in the study could be explained by TTRACE model, which claims that tone and segment are processed as a unified processing unit.

Other studies using a mismatch negativity additivity approach also found the integration between tone and segment processing. The mismatch negativity additivity approach has been proven to be a valid way for testing the integration or independence between two feature dimensions in electrophysiology studies (Choi & Tong, 2016; Gao et al., 2012; Paavilainen, Valppu, & Näätänen, 2001). To use this approach, two single deviants (varying in single dimension) and a double deviant (varying in double dimensions) are needed. If the two feature dimensions were separately processed by different neuronal populations, the sum of the mismatch negativity elicited by each single deviant would be equal to the mismatch negativity elicited by the double deviant. On the other hand, if the two dimensions are integrally processed with each other by the same neuronal population, the mismatch negativity elicited by the double deviant would not be additive to the mismatch negativity elicited by each single deviant.

Using the mismatch negativity additivity approach, Tong et al (2016) found that tone and segment are integrally processed for Cantonese speakers. In her study, three types of mismatch negativities were elicited by varying tone (e.g., /u1/ - /u3/), vowel (e.g., /u1/ - /i1/) and tone-plus-vowel (e.g., /u1/ - /i3/) of the standards (e.g., /u1/) in a passive oddball paradigm. Since the mismatch negativity elicited by tone-plus-vowel deviant was not-additive by the mismatch negativity of each single deviant, it indicated the integration between tone and vowel processing at pre-attentive stage, which supported the integrated view of TTRACE model.

While most experiments were conducted with monosyllable word perception, tone and segment processing were also tested in one study that examined sentence comprehension in Cantonese. In an event-related potential (ERP) study conducted by Schirmer et al (2005), the participants were presented with semantic congruous sentences or incongruous sentences constructed by changing either tone (e.g., bou6; "step"), rime (e.g., beng2; "biscuit") or whole syllable (e.g., gwai3; "season") of the target syllable (e.g., beng6; "illness") which appeared in the middle of the congruous sentences. Their task was to judge whether the sentence they listened to was semantic congruous or not. The ERP index concerned in the study was N400, a Negativity that peaks approximately 400 ms following target word onset over central-parietal scalp distribution, and is considered to be an index of semantic processing (Brouwer & Crocker, 2017; Hu et al., 2012; Shuai & Gong, 2014). The results showed that the rime violation and the tone violation condition elicited similar N400 component. This indicated that segment (i.e., rime in this study) and tone could be activated simultaneously and be equally important for word meaning retrieval in the context of Cantonese sentence comprehension. The results could be explained by the integrated view of TTRACE model. However, it couldn't rule out the possibility that tone and rime were still separately processed and represented as independent units but just had same weight during Cantonese sentence comprehension. Therefore, integrated vs. independent processing of tone and rime in sentence context deserved to be further investigated. To the best of my knowledge, this is the only study that examined tone and rime processing in a Cantonese sentence context. Whether the findings could be replicated and generalized to Mandarin Chinese remained unclear.

To summarize, the previous studies have not reached a consensus on the integration or independence between tone and segment processing, and little attention has been paid to tone

and segment processing in sentence level. However, in daily life conversation, a spoken word is normally embodied in a sentence rather than presented in isolation. A complete understanding of how the two sources of information together contribute to speech comprehension cannot be achieved without uncovering the integration or independence of their processing in sentence context. This study aims to provide additional evidence to the issue.

Chapter 2

The present study

2.1 Overview of the present study

Since daily speech involves sentences more than isolated words, sentence context should be a natural context for studying tone and segment processing. The present study aimed to uncover the integration or independence of tone and segment (i.e., rime) processing in a context of Mandarin sentence comprehension. Participants in the present study listened to a series of congruous and incongruous sentences and were required to judge whether each sentence was semantically congruous or not after its offset. The incongruous sentence was created by mismatching either tone (e.g., /zhong4-zhong3/), rime (e.g., /zhong4-zhuan4/) or tone-plus-rime (e.g., /zhong4-zhuan3/) in the second syllable of a word (e.g., /guan1-zhong4/; "audience") embodied in the congruous sentence. Note that tone and segment violation in this study were both achieved in rime to ensure that the violation could be detected at the same time in both condition (because tone was mainly carried by the vowel). By comparing the ERP pattern of the three violation conditions, it is possible to uncover the integration or independence of tone and rime processing. The following paragraphs would provide detailed explanations.

Since any types of mismatch would transform the original word into a pseudoword and lead to difficulty in sentence comprehension, the three violation conditions should evoke larger N400 relative to the original condition. Besides N400, another ERP component concerned in this study is P600, a positive component associated with repairing incongruous information based on the context. It was proposed that larger P600 represents easier recovering of mismatched information after incongruity is detected (Hu et al., 2012). It was hypothesized that a larger P600 would be

evoked in the three violation conditions relative to the original condition (Gouvea, Phillips, Kazanina, & Poeppel, 2010).

Hypothesis 1:

Tone and rime are integrally processed as a unified unit during Mandarin Chinese sentence comprehension.

Predicted results:

The unified unit would produce incongruity no matter the violation is in tone, rime or tone-plus-rime. Thus, it is predicted that the three violation conditions would elicit similar N400 effects. Meanwhile, the repairing of the tone, rime and tone-plus-rime mismatched syllable would all correspond to the recovering of the unified unit, leading to comparable P600 effects among the three violation conditions.

Hypothesis 2:

Tone and rime are independently processed as different units during Mandarin sentence comprehension.

Predicted results:

Tone and rime could play different roles during Mandarin sentence comprehension. As some literature showed a more important role of segment in constraining word meanings than tone (Cutler & Chen, 1997; Hu et al., 2012; Y. Tong, Francis, & Gandour, 2008; Wiener & Turnbull, 2016), incongruous segmental information should lead to more difficulties in semantic processing than incongruous tonal information. Thus, the N400 effects evoked by rime violation

should be larger than tone violation. And the N400 effects elicited by double violation should be larger than rime because tonal information is also missing in the double violation condition.

Although it is possible that tone and rime have same weight in sentence context resulting in similar N400 effects between the tone and rime violation condition as Schirmer's study (2005), the double violation condition here should still evoke stronger N400 effects compared to each single violation condition because of mismatching both tone and rime units. Therefore, differences could be found among the three violation conditions as long as tone and rime are independently processed. This ensured a stricter examination of integrated vs. independent processing of tone and rime compared with Schirmer's study, which did not have the tone-plus-rime violation condition but a complete mismatch condition violating the whole syllable.

In addition, only the tone/rime unit needs to be recovered to repair the tone/rime mismatched syllable, while both units should be recovered to repair the tone-plus-rime mismatched syllable. Different recovering tasks among the three violation conditions should result in different P600 effects. Since previous studies reported the easier recovering of tone than rime in idiom and poem contexts (Hu et al., 2012; Li et al., 2014), it was hypothesized that the P600 evoked by the tone violation condition would be larger than the rime violation condition in this study. The P600 evoked by the rime violation condition would be in turn larger than the double violation condition which additionally has a tone unit to be recovered.

2.2 Methods

2.2.1 Participants

Sixty right-handed native speakers of Mandarin Chinese (19 males, mean age: 24.8 years) were recruited in the experiments. They are mainland Chinese residents who come to

Hong Kong for postgraduate studies. All of them were paid 200 HK dollars after participation. Informed consent form was obtained before the experiment. To verify their proficiency in Mandarin Chinese comprehension, a practice session was conducted prior to the formal experiment, in which the participants heard a series of sentences highly similar to the sentences used in the formal experiment (with the same manipulations in rime and tone in the critical syllable) and did the same task (judge whether each sentence makes sentence). All of the participants completed the practice session with high accuracy, indicating that they are highly proficient in Mandarin sentence processing.

2.2.2 Stimuli

The critical materials consisted of 52 sentence frames such as “艺术团演员精彩的表演让台下的____赞口不绝” (The wonderful performance of the art troupe wins the applause of the ____ under the stage). Four types of two-syllable target were embodied in the blank of each sentence frame, including the original word (e.g., “观众”/guan1-zhong4/; “audience”) and three pseudo words made by mismatching tone (e.g., “观肿”/guan1-zhong3/), rime (e.g., “观赚”/guan1-zhuan4/) or tone-plus-rime mismatch (e.g., “观转”/guan1-zhuan3/) in the second syllable of the original word. Only the sentence frame combining with the original word would be a semantically congruous sentence, while combing with any types of pseudo word would be incongruous. In this way, 52 congruous and 156 incongruous sentences were created. The length of each sentence varied from 17 characters to 23 characters. To ensure the participants could not anticipate where the target word was and process it more naturally, the position of the target word varied across sentences. However, its position was several syllables

(at least four syllables) away from the end of the sentence to minimize the contamination of motor responses due to button pressing for congruence judgement.

By omitting the second syllable of the target word in each critical sentence (e.g., "艺术团演员精彩的表演让台下的观__赞口不绝"), a cloze probability test of 52 incomplete sentences was created and was given to 31 native Chinese college students (in mainland China), who was asked to complete the sentence and make it semantically congruous with the first suitable syllable that came to their mind. The mean cloze probability of the 52 sentences with the selected congruous syllables was above 80%. The cloze probability of the sentences completed by the three mismatch conditions was 0.

Homophone density, character frequency of the critical syllable under the four conditions were obtained from the database by Liu et al (2007). Phonological frequency was obtained from a recent database by Sun et al (2018). Analysis of variance (ANOVA) across items showed no significant difference among the four conditions in both homophone density and logarithmic transformation phonological frequency (*p*s > .10). These lexical properties of the critical syllable were summarized in Table 1.

Table 1

Sample stimuli and lexical properties across conditions

Sample sentence frame: 艺术团演员精彩的表演让台下的观__赞口不绝 (The wonderful performance of the art troupe wins the applause of the ____ under the stage)				
	Original	Tone	Rime	Double
Sample syllable	众	肿	赚	转
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. 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 syllable (*p*s < .01), which did not differ among themselves (*p*s > .10); Standard deviation (SD) is put in the parentheses.

Each participant was presented with 52 critical sentences, 13 for each condition. Their assignment was counterbalanced across and within each condition so that each sentence frame was used only once for each subject (e.g., subject 1 listened to the congruous sentence, subject 2 listened to the incongruous sentence with tone violation, subject 3 listened to the incongruous sentence with vowel violation, subject 4 listened to the incongruous sentence with tone-plus-rime violation). To balance the proportion of congruous and incongruous sentences, another 26 congruous sentence having similar length with the critical sentences were created as fillers. They were not included in the analysis. The 78 sentences in total were presented in 3 blocks. The presentation order within and across blocks were randomized across subjects.

All sentences (including 52 congruous sentence, 156 incongruous sentences and 26 filler sentences) were produced by a female native Mandarin speaker in a natural speed. During sound recording, each sentence was repeated for three times, and the one that sounded most naturally and clearly was chosen. The three types of mismatch syllable were cut from the incongruous sentences and then cross-spliced to the congruous sentence by replacing the second syllable of the original word. This ensured that the four conditions shared the same sentence frame while only differed in the second syllable (manipulated syllable) of the target word. If the four conditions elicit different ERP effects, the different ERP patterns could only be attributed to the different manipulations (Original vs. Tone vs. Rime vs. Double) on the second syllable of the target word.

2.2.3 Procedure

The participants were seated individually in a quiet room. In front of the seat, there was a computer screen and two right-and-left symmetrically placed speakers. The two speakers presented the auditory sentences bilaterally. A fixation cross preceded the auditory sentence presentation by 500 ms and remained on the screen during the entire sentence presentation. Three-hundred ms after sentence offset, a question mark "?" would appear on the screen to probe the participants to make response. The participants were informed to judge whether the sentence they heard was semantically congruous or not. Half of them were told to press "F" key for the congruous sentences and "J" for incongruous sentences, while the other half adopted the reversed pairing. They were required to keep their eyes open and to relax to reduce the likelihood of involuntary eye movements during the sentence presentation. After the participant made the judgement, the question mark disappeared, and the experiment automatically proceeded to the next trial after an interval of 2500 ms. The duration of each trial was around 12 seconds. Before the formal experiment, 12 practice sentences were presented to familiarize the participants with the task. There was a 2-minute break after each block, and the whole experiment (including EEG Setup) lasted around one and half hour.

2.2.4 EEG Recordings

The electroencephalogram (EEG) was recorded through a 64-channel cap (waveguard; ANT Neuro) placed according to the extended international 10-20 system. Cpz electrode was used as online reference, and the signal was re-referenced to the electrode placed on the nose tip in the offline analysis. Ground electrode was inlaid in the cap between Fpz and Fz electrodes. Horizontal EOG was monitored by electrodes placed on the outer canthus of each eye while 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.

2.2.5 Data Analysis

The obtained EEG data was first band pass filtered (0.05 to 30 Hz). Subsequently, the data was segmented into epochs of -200 to 1000 ms relative to the manipulated syllable (the second syllable of the target word) onset, with a baseline of 200 ms prior to the critical syllable onset. The incorrect trials or trials with amplitudes exceeding ± 100 μv were excluded from analysis. Data from 19 participants (6 males) were removed because of excessive rejection (more than 25%). The use of 25% bad trials as rejection criteria was consistent with a recent ERP study (Gaspar, Christie, Prime, Jolicœur, & McDonald, 2016). Although the removal of 19 participants might decrease the statistical power, it ensured the date quality, as the signal-to-noise ratio in these participants is relatively low due to less trails left after rejection. For the remaining 41 participants, the average number and proportion of trials maintained in the original, tone, rime and double violation conditions were 11.6 (89%), 10 (77%), 11.3 (87%) and 11.2 (86%) respectively.

Data obtained in the experiment was analyzed with JASP, a free open source platform for statistical analysis. For behavioral data, only error rate was analyzed because the delayed response task (i.e., participants withhold responses until being probed) used in this study had made reaction time uninformative. Repeated Measures Analysis of Variance (ANOVA) was conducted with Condition (original vs. tone violation vs. rime violation vs. double violation) as a within-subject independent variable. For ERP waveforms, mean signal amplitudes were computed over a series of 50 ms time windows from 0 to 1000 ms after target syllable onset. A

separate Repeated Measures ANOVA was conducted in each time window, with Condition, Anteriority (anterior vs. central vs. posterior) and Hemisphere (left vs. medial vs. right) as within-subject independent variables. For the variables Anteriority and Hemisphere, electrodes were grouped as follows: left frontal (F1, F3, F5, FC1, FC3, FC5), medial frontal (Fz), right frontal (F2, F4, F6, FC2, FC4, FC6), left central (C1, C3, C5, CP1, CP3, CP5), medial central (Cz), right central (C2, C4, C6, CP2, CP4, CP6), left parietal (P1, P3, P5, PO3, PO5, PO7), medial parietal (Pz), and right parietal (P2, P4, P6, PO4, PO6, PO8). The grouping of the electrodes was largely based on a previous language study in N400 and P600 effects (Zou, Tsang, & Wu, 2019) except the middle electrodes (Fz, Cz, Pz), which were not included in that study. The inclusion of the middle electrodes in the present study is to be consistent with a previous study that compared tone and rime processing in Mandarin idiom context (Hu et al, 2012). The multiple electrodes were then averaged and represented by one amplitude value for each condition within each time window in the following analyses. The inclusion of Anteriority and Hemisphere not only revealed the scalp distribution of the ERP signals under each condition but also improved statistical power. However, given the low spatial resolution of ERP signals, the interactions with these factors might provide little meaningful information when the primary interest was the condition main effect (Luck, 2014). Therefore, the interactions would not be discussed in details unless they were theoretically meaningful or when the observed scalp distributions contradicted with previous studies. Post-hoc comparisons were conducted with Bonferroni correction to control for inflation of Type I error. Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) was applied whenever the assumption of sphericity was violated (uncorrected *df* and corrected *p* values were reported).

Chapter 3

Results

3.1 Behavioral Data

Repeated Measures ANOVA indicated a significant effect of condition on error rates ($F(3, 120) = 15.5$, $p < .001$, $\eta_p^2 = .28$). Post hoc comparisons revealed that tone violation elicited more errors (16.1%) than the other three conditions (4.3% for original, $t(40) = 5.88$, $p < .001$, Cohen's $d = .92$; 8.8% for rime violation, $t(40) = 3.94$, $p = .002$, Cohen's $d = .61$; 7.1% for double violation, $t(40) = 4.62$, $p < .001$, Cohen's $d = .72$), which did not differ among themselves (all ps > .1).

3.2 ERP Data: 50 ms time windows

As shown in Figure 1, as compared with the original condition, the three violation conditions elicited more negative-going potentials from around 250 to 450 ms and positive-going potentials from around 450 ms to 1000 ms. Results in each 50 ms time window starting from 250 ms were reported. The pattern of condition effects could provide details about the time course of tone and rime processing. Results of the analyses are summarized in Table 2.

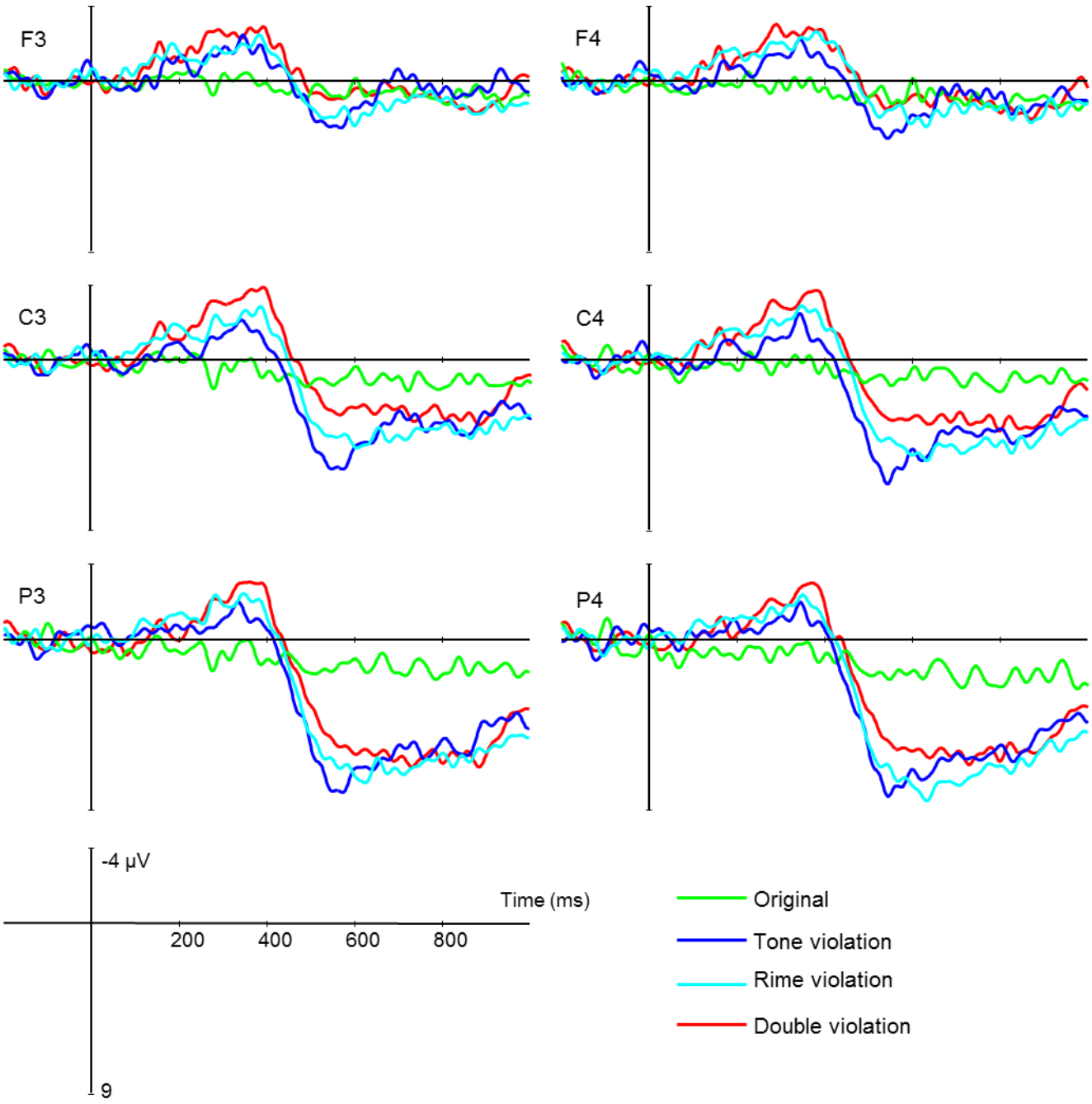

Figure 1. The ERP wave of each condition.

Table 2

F Values of the Condition × Hemisphere × Anteriority ANOVA

Time windows (msec)	Condition	Condition x Hemisphere	Condition x Anteriority	Condition x Anteriority x Hemisphere
0-50	0.850	0.975	1.883	0.985
50-100	1.042	0.478	0.766	0.532
100-150	1.301	0.453	1.466	1.007
150-200	2.549	0.720	0.500	0.942
200-250	1.747	0.320	0.923	0.815
250-300	5.784***	0.370	0.316	2.058*
300-350	5.043**	2.272*	0.316	2.394*
350-400	5.869***	3.895***	0.808	2.882**
400-450	2.011	5.878***	1.611	1.972*
450-500	1.843	7.385***	5.313**	1.312
500-550	6.946***	10.976***	6.597***	0.855
550-600	7.948***	5.691***	13.112***	0.736
600-650	7.447***	6.168***	15.093***	0.778
650-700	4.912**	4.451***	18.488***	1.108
700-750	3.838*	5.783***	16.922***	0.580
750-800	2.688*	3.004**	13.215***	0.461
800-850	2.968*	3.927***	10.896***	0.524
850-900	2.836*	3.496**	7.613***	0.608
900-950	1.714	2.550*	8.130***	0.788
950-1000	1.150	1.622	5.440***	0.434

*$p < .05$.

**$p < .01$.

*** $p < .001$.

250-300 ms time window.

There was a significant effect of Condition ($F(3, 120) = 5.78$, $p < .001$, $\eta_p^2 = .27$). As observed in Figure 1, the three violation conditions elicited larger negative potentials than the original condition (all *p*s < .001). The negativities were also significantly larger for double violation than tone violation ($t(40) = 4.81$, $p < .001$, Cohen's $d = .75$). The effect was

intermediate for rime violation, which was marginally smaller than double violation ($t(40) = 2.47$, $p = .085$, Cohen's $d = .39$) and marginally larger than tone violation ($t(40) = 2.48$, $p = .082$, Cohen's $d = .39$).

There was also a significant three-way interaction among Condition, Anteriority, and Hemisphere ($F(12, 480) = 2.06$, $p < .05$, $\eta_p^2 = .049$). Further analyses revealed that the Condition main effect was significant in all electrode sites (all ps < .01). The interaction was attributed to a larger negativity in medial electrodes than right hemisphere electrodes in the double violation condition ($t(40) = 2.59$, $p < .05$, Cohen's $d = .41$).

300-350 ms time window.

The results were mostly similar to the 250-300 ms time window. There was a significant effect of Condition ($F(3, 120) = 5.04$, $p < .01$, $\eta_p^2 = .11$). Again, the three violation conditions elicited larger negativities than the original condition (all ps < .001). The negativities were larger for double violation than tone violation ($t(40) = 3.21$, $p < .01$, Cohen's $d = .50$). The effect was intermediate for rime violation, although it was not significantly different from double violation and tone violation (ps > .1).

The three-way interaction among Condition, Anteriority, and Hemisphere was significant ($F(12, 480) = 2.39$, $p < .05$, $\eta_p^2 = .056$). The interaction could be attributed to smaller negativities in medial electrodes than left hemisphere electrodes in the three violation conditions. The negativities were also smaller in medial electrodes than right hemisphere electrodes in the rime and double violation conditions (all ps < .05). There was also a two-way interaction between Anteriority and Hemisphere in the double violation condition ($F(4, 160) = 6.34$, $p < .001$, η_p^2

= .14), because the above hemispheric differences were limited to anterior and central electrodes (all *p*s < .05).

350-400 ms time window.

Similar to previous time windows, there was a significant main effect of Condition ($F(3, 120) = 5.87$, $p < .001$, $\eta_p^2 = .13$). The three violation conditions all evoked larger negativities compared with the original condition (all *p*s < .001). Besides, the negativities elicited by the double violation condition were larger than the rime violation condition ($t(40) = 2.72$, $p < .05$, $d = .43$), which were in turn larger than the tone violation condition ($t(40) = 3.80$, $p = .001$, Cohen's $d = .59$).

In addition to the main effect of condition, there was also a two-way interaction between Condition and Hemisphere ($F(6, 240) = 3.90$, $p < .001$, $\eta_p^2 = .089$) and a three-way interaction among Condition, Hemisphere and Anteriority ($F(12, 480) = 2.88$, $p < .005$, $\eta_p^2 = .067$). The two-way interaction between Condition and Hemisphere was due to smaller negativities in medial electrodes than left and right electrodes in all violation conditions. The negativities were also smaller in right hemisphere electrodes compared with left hemisphere electrodes only in the tone violation condition (*p*s < .05). The three-way interaction was significant because the above hemisphere effects were only limited in anterior and central electrodes for the tone and double violation conditions (all *p*s < .05).

400-450 ms time window.

Unlike the previous time windows, the Condition main effect was not significant. However, Condition interacted with Hemisphere ($F(6, 240) = 5.88$, $p < .001$, $\eta_p^2 = .13$). There was also a three way-interaction among Condition, Hemisphere and Anteriority ($F(12, 480) = 1.97$, $p < .05$,

η_p^2 = .047). Further analyses indicated that Condition had an effect only in left hemisphere electrodes (F(3, 120) = 2.74, p < .05, η_p^2 = .064), such that the rime and double violation conditions evoked larger negativities than the original condition (ps < .01). No differences were found between the tone violation and the original conditions or among the tone, rime, and double violation conditions (all ps > .05).

The two-way interaction between Condition and Hemisphere also indicated that there were smaller negativities in medial electrodes than left and right hemisphere electrodes for the three violation conditions and smaller negativities in right than left hemisphere electrodes in the tone violation condition. The three-way interaction indicated that the above Hemisphere effects were only limited to anterior and central electrodes (ps < .01).

450-500 ms time window.

The Condition main effect was also not significant. Yet, the two-way interactions between Condition and Hemisphere (F(6, 240) = 7.39, p < .001, η_p^2 = .16) and between Condition and Anteriority (F(6, 240) = 5.31, p < .005, η_p^2 = .12) were significant. Further analyses indicated that Condition had an effect in medial electrodes (F(3, 120) = 3.24, p < .05, η_p^2 = .08), such that both the tone and rime violation conditions evoked larger positivities than the original condition (all ps < .05). Tone violation also elicited larger positivities than the double violation condition (t(40) = 2.99, p < .05, Cohen's d = .47). Other Condition differences were non-significant (all ps > .1).

The interaction between Condition and Hemisphere indicated that there were larger positivities in medial electrodes than left and right hemisphere electrodes in the three violation conditions. The positivities were also larger in the right than left hemisphere in the tone violation

condition (all *p*s < .05). The interaction between Condition and Anteriority indicated that there were larger positivities in central and posterior electrodes than anterior electrodes in the three violation conditions. The positivities were also larger in posterior electrodes than central electrodes for the rime and double violation conditions (all *p*s < .05).

500-550 ms time window.

From this time window to the 850-900 ms time window, the patterns of results were highly similar. The main effect of Condition became significant again ($F(3, 120) = 6.95$, $p < .001$, $\eta_p^2 = .15$). The three violation conditions all evoked larger positivities than the original condition (*ps* < .001). The positivities in the tone violation condition were also larger than the rime condition ($t(40) = 3.92$, $p < .001$, Cohen's $d = .61$), which were in turn larger than the double violation condition ($t(40) = 3.33$, $p < .01$, Cohen's $d = .52$).

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 11.0$, $p < .001$, $\eta_p^2 = .22$). The interaction was attributed to larger positivities in the medial electrodes than left and right hemisphere electrodes for the three violation conditions. Right hemisphere electrodes also had larger positivities than left hemisphere electrodes in both the tone and rime violation conditions (*ps* < .01), but not in the double violation condition.

The two-way interaction between Condition and Anteriority was significant ($F(6, 240) = 6.60$, $p < .001$, $\eta_p^2 = .14$). The interaction could be attributed to larger positivities in posterior electrodes than central electrodes, which were in turn larger than anterior electrodes in all three violation conditions (*ps* < .001).

550-600 ms time window.

The main effect of Condition was significant ($F(3, 120) = 7.95$, $p < .001$, $\eta_p^2 = .17$). The three violation conditions all evoked larger positivities than the original condition (*ps* < .001). In addition, the positivities evoked by the tone violation condition were larger than the rime violation condition ($t(40) = 2.84$, $p < .05$, Cohen's $d = .44$), which were in turn larger than the double violation condition ($t(40) = 3.60$, $p < .005$, Cohen's $d = .56$).

There was a significant two-way interaction between Condition and Hemisphere ($F(6, 240) = 5.69$, $p < .001$, $\eta_p^2 = .13$). The interaction was attributable to larger positivities in medial electrodes than left and right hemisphere electrodes in the three violation conditions. The positivities in right hemisphere electrodes were also larger than left hemisphere electrodes in both the tone and rime violation conditions (*ps* < .05), but not in the double violation condition.

The two-way interaction between Condition and Anteriority was also significant ($F(6, 240) = 13.1$, $p < .001$, $\eta_p^2 = .25$). The positivities in posterior electrodes were lager than central electrodes, which were in turn larger than anterior electrodes in the three violation conditions (*ps* < .001).

600-650 ms time window.

The main effect of Condition was significant ($F(3, 120) = 7.45$, $p < .001$, $\eta_p^2 = .16$). The three violation conditions all evoked larger positivities than the original condition (*ps* < .001). Besides, both tone and rime violation conditions elicited larger positivities than the double violation condition $t(40) = 3.53$, $p < .005$, Cohen's $d = .55$ and $t(40) = 4.98$, $p < .001$, Cohen's $d = .78$, respectively). The two conditions did not differ between themselves ($p > .1$).

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 6.17$, $p < .001$, $\eta_p^2 = .13$), which could be attributed to larger positivities in medial and right than

left hemisphere electrodes in the tone and rime violation conditions (*ps* < .001) and larger positivities in medial than left and right hemisphere electrodes (*ps* < .05).

Although the two-way interaction between Condition and Anteriority was significant ($F(6, 240) = 15.1$, $p < .001$, $\eta_p^2 = .27$), the same pattern of Anteriority effect was found in all conditions. Specifically, posterior electrodes had larger positivities than central electrodes, which were in turn larger than anterior electrodes in all conditions (*ps* < .05).

650-700 ms time window.

The main effect of Condition ($F(3, 120) = 4.91$, $p < .005$, $\eta_p^2 = .11$) was significant. The three violation conditions all evoked larger positivities than the original condition (*ps* < .001). There were larger positivities in the rime than the double violation conditions ($t(40) = 4.08$, $p < .001$, Cohen's $d = .64$). However, the tone violation condition did not differ from the rime and double violation conditions (*p*s > .1).

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 4.45$, $p < .001$, $\eta_p^2 = .10$), which could be attributed to larger positivities in medial and right hemisphere electrodes than left hemisphere electrodes only in the tone and rime violation conditions (*ps* < .001). In the rime violation condition, the positivities were also larger in medial than right hemisphere electrodes ($p < .05$).

The two-way interaction between Condition and Anteriority was significant ($F(6, 240) = 18.5$, $p < .001$, $\eta_p^2 = .32$). In the original condition, posterior electrodes had significantly larger positivities than central electrodes ($p < .05$), which were in turn marginally larger than anterior electrodes ($p = .09$). In the three violation conditions, posterior electrodes had significantly larger positivities than central electrodes, which were in turn larger than anterior electrodes (*p*s < .05)

700-750 ms time window.

The main effect of Condition was significant ($F(3, 120) = 3.84$, $p < .05$, $\eta_p^2 = .088$). The pattern of Condition effect was exactly identical to the 650-700 ms time window. The three violation conditions all evoked larger positivities than the original condition (*ps* < .001). There were larger positivities in the rime than the double violation conditions ($t(40) = 2.80$, $p < .05$, Cohen's $d = .44$). However, the tone violation condition did not differ from the rime and double violation conditions (*p*s > .05).

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 5.78$, $p < .001$, $\eta_p^2 = .13$), which was attributable to larger positivities in medial and right hemisphere electrodes than left hemisphere electrodes only in the tone and rime violation conditions (*ps* < .001). No differences were found between medial and right hemisphere electrodes (*p* > .05).

Although the two-way interaction between Condition and Anteriority was significant ($F(6, 240) = 16.9$, $p < .001$, $\eta_p^2 = .30$), the same pattern of Anteriority effect was found in all conditions. Specifically, posterior electrodes had larger positivities than central electrodes, which were in turn larger than anterior electrodes in all conditions (*ps* < .05).

750-800 ms time window.

The main effect of Condition was significant ($F(3, 120) = 2.69$ $p = .05$, $\eta_p^2 = .063$). The three violation conditions all evoked larger positivities than the original condition (*ps* < .001), while no differences were found among the three violation conditions (*ps* > .10).

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 3.00$, $p < .01$, $\eta_p^2 = .07$), which was attributable to larger positivities in medial and right hemisphere electrodes than left hemisphere electrodes only in the tone and rime violation conditions (*ps* < .01). No differences were found between medial and right hemisphere electrodes ($p > .1$).

Although the two-way interaction between Condition and Anteriority was significant ($F(6, 240) = 13.2$, $p < .001$, $\eta_p^2 = .25$), the same pattern of Anteriority effect was found in all conditions. Specifically, posterior electrodes had larger positivities than central electrodes, which were in turn larger than anterior electrodes in all conditions (*ps* < .05).

800-850 ms time window.

The main effect of Condition was significant ($F(3, 120) = 2.97$, $p < .05$, $\eta_p^2 = .069$). The three violation conditions all evoked larger positivities than the original condition (*ps* < .001), while no differences were found among the three violation conditions (*ps* > .10).

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 3.93$, $p < .001$, $\eta_p^2 = .089$), which could be attributed to larger positivities in medial and right hemisphere electrodes than left hemisphere electrodes only in the tone and rime violation conditions (*ps* < .001). No differences were found between medial and right hemisphere electrodes ($p > .1$).

The two-way interaction between Condition and Anteriority was significant ($F(6, 240) = 10.9$, $p < .001$, $\eta_p^2 = .12$). There were no Anteriority effects in the original condition. In the three violation conditions, posterior electrodes had significantly larger positivities than central electrodes, which were in turn larger than anterior electrodes (*p*s < .001).

850-900 ms time window.

The main effect of Condition was significant ($F(3, 120) = 2.84$, $p = .05$, $\eta_p^2 = .066$). The three violation conditions all evoked larger positivities than the original condition (*ps* < .001), while no differences were found among the three violation conditions (*ps* > .10).

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 3.93$, $p < .001$, $\eta_p^2 = .089$), which could be attributed to larger positivities in medial and right hemisphere electrodes than left hemisphere electrodes only in the rime violation condition (*ps* < .001). No differences were found between medial and right hemisphere electrodes (*p* > .1).

The two-way interaction between Condition and Anteriority was significant ($F(6, 240) = 7.61$, $p < .001$, $\eta_p^2 = .16$). In the original condition, posterior had larger positivities than Central and anterior electrodes (*ps* < .01). In the three violation conditions, central electrodes also had larger positivities than anterior electrodes (*ps* < .001).

900-950 ms time window.

Unlike the previous time windows, the Condition main effect was not significant. However, Condition interacted with Hemisphere ($F(6, 240) = 2.55$, $p < .05$, $\eta_p^2 = .060$) and Anteriority ($F(6, 240) = 8.13$, $p < .001$, $\eta_p^2 = .17$). Further analyses indicated that Condition had an effect only in posterior electrodes ($F(3, 120) = 3.85$, $p < .05$, $\eta_p^2 = .088$), such that all three violation conditions evoked larger positivities than the original condition (*ps* < .01). No differences were found among the three violation conditions (*ps* > .10).

The two-way interaction between Condition and Hemisphere also indicated that there were Hemisphere effects only in the tone and rime violation conditions. Medical and right hemisphere

electrodes had larger positivities than left hemisphere electrodes (*ps* < .01). No differences were found between medial and right hemisphere electrodes (*p* > .1).

The two-way interaction between Condition and Anteriority also indicated that there were Anteriority effects only in the three violation conditions. Posterior electrodes had larger positivities than central electrodes, which were in turn larger than anterior electrodes (*ps* < .001).

950-1000 ms time window.

As in the 900-1000 ms time window, the Condition main effect was not significant. Only the two-way interaction between Condition and Anteriority ($F(6, 240) = 5.44$, $p < .001$, $\eta_p^2 = .12$) reached significance. Further analyses indicated that the Condition effects were not significant in any electrode regions (*ps* > .10). The two-way interaction also indicated that in the original condition, posterior electrodes had larger positivities than central and anterior electrodes (*ps* < .01). On the other hand, in the three violation conditions, posterior electrodes had larger positivities than central electrodes, which were in turn larger than anterior electrodes (*ps* < .01).

To summarize, the successive time windows analyses showed larger early negativities in 250-450 ms and larger positivities in 450-950 ms produced by the violation conditions compared with the original condition. The negativities evoked by double violation were larger than rime violation, which were in turn larger than tone violation. The positivities elicited by tone violation were larger than rime violation in 500-600 ms, but they were similar to each other in 600-950 ms. The positivities evoked by tone violation and rime violation were both larger than double violation in 500-650 ms, while only rime violation continued to be larger than double violation in 650-750 ms. The positivities elicited by the three violation conditions were comparable to each other in 750-950 ms. The results were summarized in the following Table 3.

Table 3

The comparisons among the four conditions across time windows

Time	Polarity	Amplitude
250~400	Negative	Double > Rime > Tone > Original
400~450	Negative	Double, Rime > Original
450~500	Positive	Tone, Rime > Original; Tone > Double
500~600	Positive	Tone > Rime > Double > Original
600~650	Positive	Tone = Rime > Double > Original
650~750	Positive	Rime > Double; Tone, Rime, Double > Original
750~950	Positive	Tone = Rime = Double > Original

3.3 ERP Data: N400 and P600

Although the 50-ms time window analyses provided important details about the time course of effects, the multiple comparisons could inflate Type-I error. Moreover, as shown above, the patterns of effects were usually similar in adjacent time windows. To complement the above analyses, broader time windows were constructed by combining successive time windows that showed significant Condition effects of similar patterns (using a cut-off p-value of .01). Accordingly, two clusters were identified. The first cluster was 250-400 ms, in which the waveforms were negative and corresponded to N400. The second cluster was 500-700 ms, in which the waveforms were positive and corresponded to P600. The scalp map of N400 and P600 was depicted in Figure 2.

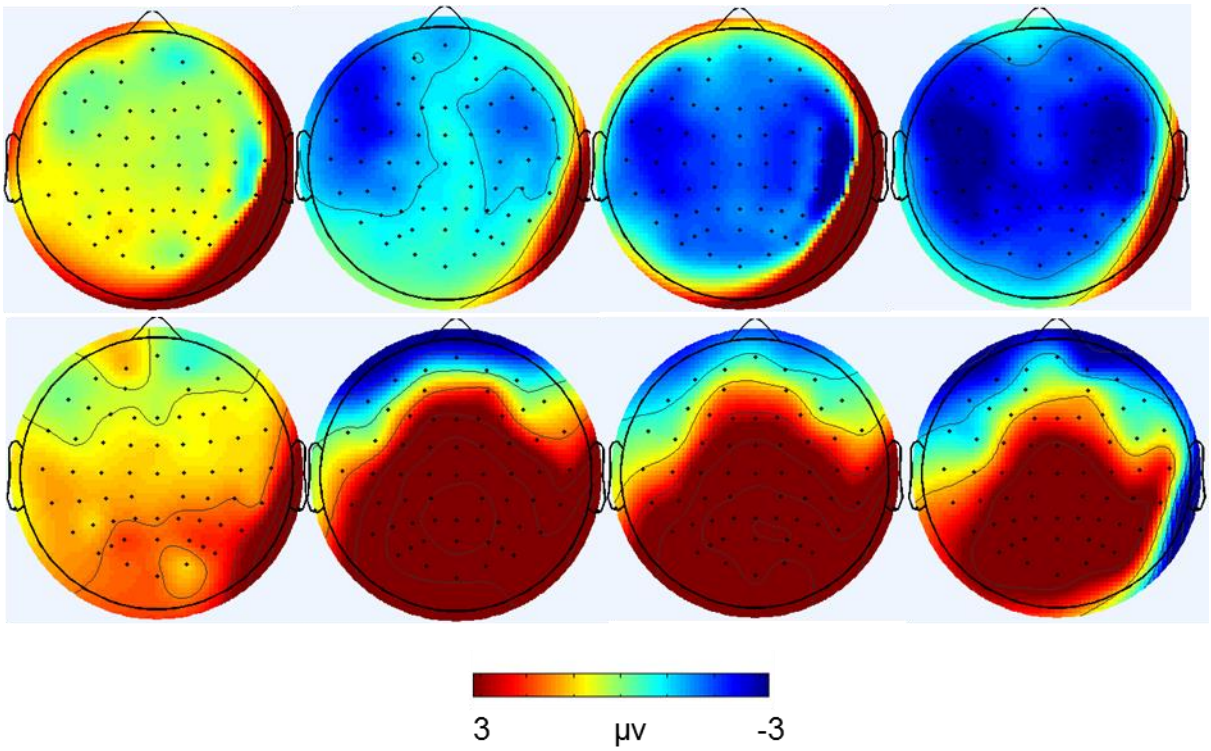

Figure 2. The Scalp map of N400 and P400 in each condition.

Note. The upper row is N400 and lower row is P600. From Left to Right is Original, Tone violation, Rime violation, and Double violation respectively.

N400

There was a significant main effect of Condition ($F(3, 120) = 6.79$, $p < .001$, $\eta_p^2 = .15$). The three violation conditions evoked larger N400 than the original condition (*ps* < .001). In particular, the double violation condition had the largest N400, which was marginally larger than the rime violation condition ($t(40) = 2.57$, $p = .063$, Cohen's $d = .40$) and was significantly larger than the tone violation condition ($t(40) = 5.30$, $p < .001$, Cohen's $d = .83$). The N400 in the rime violation condition was also significantly larger than the tone violation condition ($t(40) = 2.83$, $p < .05$, Cohen's $d = .44$).

The three-way interaction among Condition, Hemisphere and Anteriority was significant ($F(12, 480) = 2.64$, $p = .01$, $\eta_p^2 = .062$). It was shown that significant Hemisphere effects could be found at all Anteriority levels in the rime violation condition. The Hemisphere effects were limited to anterior and central regions in the double violation condition and to the anterior region in the tone violation condition (*ps* < .05). Further analyses showed that in the rime violation and double violation conditions, left and right hemisphere electrodes had larger N400 than medial electrodes. In the tone violation condition, left hemisphere electrodes had significantly larger N400 than medial electrodes ($p < .001$), but the difference between right hemisphere and medial electrodes was only marginal ($p = .056$).

P600

There was a significant main effect of Condition ($F(3, 120) = 8.19$, $p < .001$, $\eta_p^2 = .17$). The three violation conditions evoked larger P600 than the original condition (*ps* < .001). While the rime and tone violation conditions did not differ between themselves ($t(40) = 1.01$, $p = 1.00$, Cohen's $d = .16$), both evoked larger P600 than the double violation condition ($t(40) = 4.53$, $p < .001$, Cohen's $d = .71$ and $t(40) = 5.01$, $p < .001$, Cohen's $d = .78$, respectively).

Although the analysis of the mean amplitude of P600 in the 500-700 ms window failed to reveal the difference between tone and rime condition, visual inspection of the signal wave (in Figure 1) suggested that the P600 elicited by tone violation was initially larger than rime violation, but was a bit smaller than rime violation in the later half of the window. Such visual inspection results were consistent with the findings in the successive time window analysis that tone violation evoked larger positivities than rime violation from 500-600 ms.

The two-way interaction between Condition and Hemisphere was significant ($F(6, 240) = 7.74$, $p < .001$, $\eta_p^2 = .16$). There were no Hemisphere effects in the original condition. In the three violation conditions, medical electrodes had larger P600 than left and right hemisphere electrodes (*ps* < .01). In both the tone and rime violation conditions, right hemisphere electrodes also had larger P600 than left hemisphere electrodes (*ps* < .001).

Althoguh the two-way interaction between Condition and Anteriority was also significant ($F(6, 240) = 14.8$, $p < .001$, $\eta_p^2 = .27$), the same pattern of Anteriority effect was found in all conditions. Specifically, posterior electrodes had larger P600 than central electrodes, which were in turn larger than anterior electrodes in all conditions (*ps* < .01).

Chapter 4

Discussion

Although variation in tone is mainly carried by rime, whether tone and rime are integrally or independently processed remained inconclusive. As most previous studies only considered tone and segment processing in isolated words, the findings may not be generalizable to daily speech comprehension, where a spoken word is processed in a natural sentence context. The present study filled this gap, and further explored the integration or independence between tone and segment processing during Mandarin sentence comprehension.

The behavioral data showed that participants made more errors for tone violated sentences compared with the original, rime violated and double violated sentences. The ERP results indicated that, compared with the original condition, all violated conditions evoked larger negativities in the 250-400 ms time window, and larger positivities in the 500-700 ms time window after the critical syllable onset. In addition, the negativities elicited by the double violation was larger than that by the tone violation, while the effect of the rime violation was intermediate between the tone and double violations. The positivities evoked by the tone violation condition were larger than the rime violation condition in the 500-600 ms time window, but they were comparable in the 600-700 ms time window. The positivities evoked by the tone and rime violations (single violation) were both larger than the double violation condition.

4.1 The Early Negative Effect

Previous studies using sentence comprehension task reported N400 effects evoked by tone and segment violations(Schirmer et al., 2005). N400 is an ERP component that represents semantic processing (Brouwer & Crocker, 2017; Hu et al., 2012). It typically has a latency of

300-500 ms after the stimuli onset. The amplitude of N400 is an index of the difficulty in semantic processing. The higher the amplitude is, the more difficulty people encounter in semantic integration. The latency of the early negative effects in the present study was 250-400 ms, which was a bit earlier than the typical N400. However, as pointed out by Van Petten (1999), information about a spoken word unfolds in time in the auditory tasks, and the processing of mismatched word could diverge from original word once the contextual mismatched information is detected, even the word identification may not yet be complete. In other words, it is speculated that the auditory task combing with the strong context might lead to the earlier N400 effects in the present study. Evidence about this speculation could be found in previous studies that showed earlier N400 effects in auditory tasks than visual tasks (Hu et al., 2012; Schirmer et al., 2005), and reported earlier N400 effect when a strong context was presented (Shuai & Gong, 2014). For example, in a study conducted by Shuai and Gong (2014), the author found an early onset N400 effect (300-350 ms) in a highly constraining context. Likewise, in a highly predictive sentence context, Schirmer et al (2005) also reported an early N400 (250-450 ms) in response to tone and segment violations. Based on these studies, the early negativities were identified as auditory N400 in the present study.

Since both tone and segment violation evoked larger N400 effects than the original condition in the present study, it again provided evidence that tonal and segmental information are both accessed during Mandarin sentence comprehension, which replicated previous findings (Brown-Schmidt & Canseco-Gonzalez, 2004; Cutler & Chen, 1997; S. Liu & Samuel, 2007; Schirmer et al., 2005). Besides, the larger N400 evoked by rime violation compared with tone violation indicated that the participants might encounter more difficulties in the semantic processing of a rime mismatched syllable than a tone mismatch one during sentence

comprehension. In other words, rime appeared to be more important in meaning activation than tone. The larger N400 elicited by double violation than single violation conditions (rime or tone) showed that it is the hardest to process meanings when both tone and rime were incongruous.

If tone and segment are integrally processed as a unified unit as hypothesis 1 assumed, the three violation conditions would lead to comparable difficulties in semantic processing, which could not explain the different N400 effects evoked in the three violation conditions. Instead, the different N400 effects among the three violation conditions supported the independence between tone and segment processing (hypothesis 2). It was speculated that the different N400 patterns were related to the independent representations of tone and rime in mental lexicon, as well as unequal information value carried by tone and rime in constraining word identity.

In Mandarin Chinese, syllables are composed of an initial consonant, a rime (vowel) as well as a lexical tone. There are 23 consonants, 39 rimes and 4 tones, which can be combined to form around 1300 unique syllable-plus-tone units (note that not all combinations exist; Y. Tong et al., 2008). As there are much more rimes (39) than tone (4), the probability of consonant-plus-rime pronounced in each tone is much higher than that of consonant-plus-tone integrated with each rime. Therefore, tone has lower information value in constraining word identity than rime. Participants might encounter more difficulty in recognizing the word when cues with higher information value were missing, resulting in a larger N400 for the rime violation condition compared with the tone violation condition. And it would be the hardest for participants to comprehend the sentences when both tone and rime mismatched, leading to the largest N400 in the double violation condition.

This information value idea was also supported by previous studies with lexical decision task. For example, Wiener and Turnbull (2016) found that tone violated pseudowords were more easily accepted as real words relative to rime violated pseudowords, showing that tone was not as efficient as rime for constraining spoken word recognition. The more weight of rime in lexical access could also be found in idiom context. For instance, Hu et al (2012) found that rime violation in the final syllable of idiom evoked larger N400 than tone violation when the participants judged the correctness of the final syllable in each idiom. In an early study conducted by Taft and Chen (1992), when the participants were required to decide whether a pair of characters are homophones or not, they experienced more difficulties in rejecting tone mismatched pairs (e.g., “曲” and “去”, i.e., /qu3/ and /qu4/) as homophones compared with rime mismatched pairs (e.g., “气” and “去”, i.e., /qi4/ and /qu4/). Together, these studies indicated the advantages of rime over tone in word identification.

It should also be noted that the N400 differences between the tone and rime violation conditions in this study were inconsistent with a previous study with Cantonese speakers, which showed concurrent and comparable N400 effects evoked by tone and rime violations (Schirmer et al., 2005). One possible reason that led to the inconsistency between the two studies is the different language teaching methods adopted in mainland China and Hong Kong. Children in mainland China are taught an alphabetic system of Chinese (Pinyin) when they learn words. As tone and segment are explicitly distinguished in pinyin learning (Wang, Li, & Lin, 2015; Shu, Peng, & McBride-Chang, 2008), and only segment was needed for typing pinyin to input character on computers or mobile phones in daily life, mainland participants might realize that segment has higher information value than tone in constraining word identity and be more sensitive to segment (Wen, Filik, & van Heuven, 2018). In contrast, children in Hong Kong are

taught orthographic-to-pronunciation mapping without the mediation of pinyin (McBride, Bialystok, Chong, & Li, 2004). Thus, Hong Kong participants might be less likely to distinguish the information value of tone and segment, while they tend to process tone and segment as a unified unit. Evidence about the effects of pinyin on tone and segment processing was found in a study conducted by Wang et al (2015) using priming naming paradigm. In the study, the author found that tone and segment were accessed as a unified unit during character naming when visual character primes were presented, while they were separately activated as independent units when pinyin primes were presented, indicating that pinyin could promote the independent activation of tone and segmental information in spoken word planning. However, it would be interesting for future studies to examine whether pinyin training really moderated the independence vs. integrated processing of rime and tone during auditory sentence comprehension.

Another possible reason is the closer link between rime and tone in Cantonese than Mandarin. Specifically, in Cantonese, the rimes that end with the three stop codas (i.e., /p/, /t/, /k/) always carry tone 1, 3 or 6 (Cutler & Chen, 1997). In other words, certain rimes often co-occur with certain tones, which might result in the integrated processing of tone and rime during Cantonese sentence comprehension. Yet, this speculation also needs to be verified in future studies.

4.2 The Late Positive Effect

The early negativity observed in this study was followed by a late positive component in the window of 500-700 ms after critical syllable onset. The positive component was larger for the three violation conditions compared with the original condition. Besides, the positive components for tone and rime violation conditions were also larger than the double violation

condition. The positive component for the tone violation condition was initially larger than the rime violation condition in the time window of 500-600 ms, but did not differ from rime violation condition in the window of 600-700 ms after stimuli onset. This positive component had maximum over posterior sites for all conditions. Give its latency and scalp distribution, the positive component was recognized as P600 in the present study. The P600 elicited by mismatched syllables was also found in previous language studies and were associated with repairing mechanism for mismatched information (Hu et al., 2012; Schirmer et al., 2005). The repairing mechanism implies that when participants found it hard to integrate the mismatched syllable into the semantic context, they might try to restore the syllable based on its acoustic feature as well as the highly predictive context. In addition to the repairing mechanism, P600 were also related to syntactic error detection and more elaborated semantic processing (Brouwer & Crocker, 2017; Gouvea et al., 2010). Based on these previously discovered properties of P600 and given that semantic processing difficulty already emerged in the earlier time windows (250-400 ms), it was proposed that the P600 effects here represented the efforts participants devoted to recover the mismatched syllable to fit into the context.

The initial larger P600 evoked by the tone violation condition compared with the rime violation might indicate a faster start of the recovering process when mismatched in tone than rime (Gouvea et al., 2010). Yet, the comparable P600 in the later window (600-700 ms) showed that equivalent mental resources were allocated to recover the mismatched rime even its repairing was not as fast as tone. A larger P600 elicited by tone violation than rime violation was also found in previous studies with idiom and poem context (Li, Wang, & Yang 2014; Hu et al., 2012). These studies together indicated the advantages of tone over rime in recovering.

If tone and segment are integrally processed as a unified unit as hypothesis 1 assumed, recovering tone and rime mismatched syllable would both mean the recovering of the whole rime-plus-tone unit and would result in similar P600 effects. However, the different P600 effects evoked by tone and rime violations contradicted with the view of integrated processing, but indicated the independence and representation between tone and segment processing during Mandarin sentence comprehension (hypothesis 2). It was further speculated that the differences between tone and rime were attributed to two factors: *Retrieval difference* and *Daily speech experience in normalizing tone*.

Retrieval difference. As aforementioned, the number of rime (39) is much higher than tone (4) in Mandarin Chinese, therefore, the probability of tone combining with each consonant-plus-rime is much lower than that of rime combing with each consonant-plus-tone. Consequently, participants might only need to try at most three alternative tones to recover the tone mismatched syllable, while they had to search a larger number of rimes to repair the rime mismatched one. This advantage in retrieval might lead to faster recovering of mismatched tone than mismatched rime, resulting in a larger P600 in the window of 500-600 ms. The retrieval difference idea was also supported by a recent behavioral study conducted by Wiener and Turnbull (2016). In the study, Mandarin speakers heard pseudowords (e.g., /su3/) and were required to use either tone change (e.g., /su4/), consonant change (e.g., /tu3/) or rime change strategies (e.g., /si3/) to make them real words. The author found that the speed and accuracy of the participants were the highest in using the tone change strategy, intermediate in using the consonant change strategy, and the lowest in using the rime change strategy. The pattern was consistent with the retrieval difference hypothesis that easier retrieval leads to faster recovery.

Daily speech experience in normalizing tone. In addition to the retrieval difference, the proficiency in recovering tone mismatched syllable may also be associated with people's daily speech experience in dealing with tone change. Tone is physically more variable because the realization of tone (pitch variation) can be easily affected by the context and individual differences (Cutler & Chen, 1997; Schirmer et al., 2005). For example, when pronouncing two-syllable word, "雨伞" (/yu3-san3/, umbrella), the third tone of the first syllable is changed into the second tone. This phenomenon is called "tone sandhi". Besides, the pitch height of the same lexical tone also varies across different forms of sentence. For instance, the same lexical tone produced in questions normally has higher pitch than in statements. To ensure smooth communication, listeners have to process the input tone more flexibly (Schirmer et al., 2005; Shao & Zhang, 2019). The long-term experience in normalizing variable tone might lead to better restoration for mismatched tone than rime in the sentence comprehension context, resulting in initially larger P600 in the tone violation condition. However, the advantages of tone in recovering also make the tone mismatched sentences more acceptable than rime and double mismatch ones. Therefore, participant made more error decisions in the tone violation condition compared with the rime and double violation condition.

As recovering the mismatched syllable also relies on the input acoustic information (Schirmer et al., 2005), missing more acoustic cues would naturally lead to larger difficulty in repairing. Thus, even the context is highly constraining in the present study, it might still be too hard for participants to recover the double violated syllables with more valid acoustic cues lost. As a result, participants might be less willing to reanalyse the double violated syllables than the single violated ones, which led to a smaller P600 in the double violation condition than each single violation condition (the tone and rime violation conditions). If tone and rime are processed

as a unified unit, the repairing of the double violated syllable should require the recovering of the same unified unit as each single violated syllable. Thus they should evoke similar P600 effects. Again, the different P600 effects between the double violation and each single violation conditions here showed that tone and rime are not processed as a unified unit, but as independent units in the context of Mandarin sentence comprehension.

4.3 Theoretical contribution

With the violation paradigm, a functional dissociation of tone and segment during Mandarin sentence comprehension was discovered. If tone and segment (rime in this study) are integrally processed as one unit as described by TTRACE model (Choi & Tong, 2016), they should play comparable roles during spoken word perception. However, the respective advantages and disadvantages of tone and rime found in this study suggested that they were not integrally processed as a unified unit, but were independently processed as different units in the context of Mandarin sentence comprehension. The independence between tone and rime processing was consistent with the view of the modified TRACE model regarding the independent mental representations and processing of tonal and segmental information (Ye & Connine, 1999).

The modified TRACE model additionally suggested an important role of context in spoken word perception. Specifically, a stronger context could provide more facilitation to tone processing than rime processing, which could dilute the disadvantages of tone in word identification. Yet, even in the highly predictive sentence context in this study, tone still showed weaker power than rime in constraining word identity, which contradicted with the context effects described by the modified TRACE model. On the other hand, context might contribute

more to the recovering of the mismatched tone, such that tone showed advantages over rime in recovering stage under the strong sentence context.

Given the respective strengths and weaknesses of tone and rime across different stages during Mandarin sentence comprehension, a dynamic model of tone and segment processing was proposed and depicted in Figure 3 and Figure 4. In this model, tone and segment are also represented as independent units and all the units are interconnected with each other in mental lexicon. However, the interconnection pattern varies across different processing stages. In the syllable identification stage, tone and segment interact with each other and contribute together for word identification. The connection between segment and syllable is stronger than that between tone and syllable because segment plays a dominant role in accessing the syllable unit. In the reanalysis stage, connections also exist between context and acoustic input information. The connection between context and format frequency (features rime) is weaker than that between context and fundamental frequency (features tone), leading to more difficult reanalysis (slower) of segment (i.e., rime in this study) than tone. Tone and segment unit are also connected with each other in the reanalysis stage. Specifically, correct segmental information is helpful for recovering tone, and recovering segment could also be facilitated by the correct tonal information. Two reanalysis paths (tone and segment reanalysis paths) could both be exploited to recover the syllable unit to fit into the context.

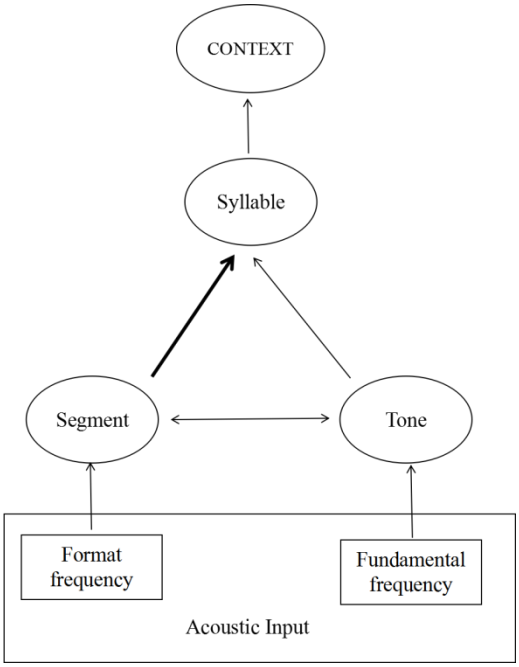

Figure 3. Dynamic model of Mandarin spoken word processing: Syllable identification stage.

Note. Thick line represents stronger connection than thin line.

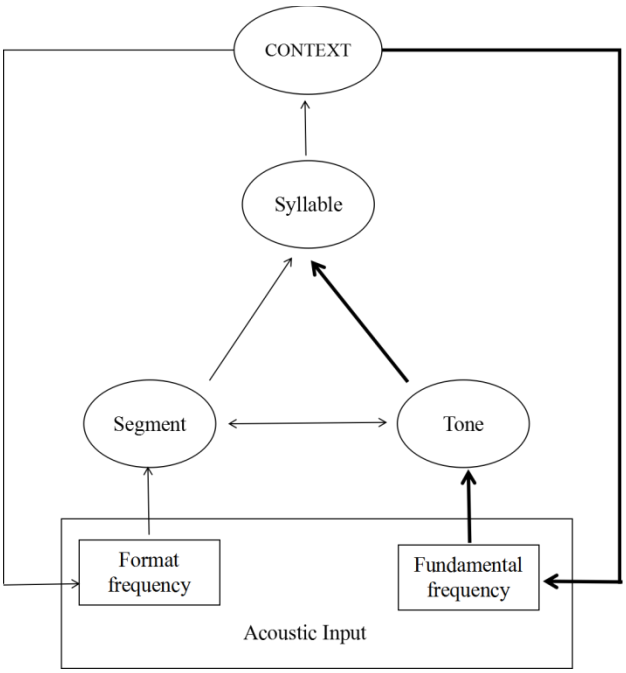

Figure 4. Dynamic model of Mandarin spoken word processing: Reanalysis stage.

Note. Thick line represents stronger connection than thin line.

4.4 Limitation

Although the present study succeeded in revealing the independence between tone and segment (rime) processing during Mandarin sentence comprehension, it also has limitation. As the four conditions (Original, Tone, Rime, Double) differed in the critical syllables embodied in the same sentence frames, the different ERP effects obtained in each condition might be partly due to the processing of different syllables, even when some important properties of these syllables, such as homophone density and phonological frequency have been matched. Future studies could adopt a within-item design, using the same syllables in different sentence frames, such that the same syllable in one sentence frame as one condition could be interchanged to the other sentence frame as other condition. When the same syllables were used across conditions, a better comparison between tone and rime processing could be obtained.

In addition, the findings of the present study were inconsistent with a previous study on Cantonese speakers, which could be explained by the integrated view of tone and rime processing (Schirmer et al., 2005). It was speculated that the discrepancies between the two studies were due to the different teaching methods adopted in mainland China and Hong Kong, as well as the varying degree of link between rime and tone in Cantonese and Mandarin. The independent processing of tone and rime found in this study was associated with pinyin teaching in Mainland China. And the integrated processing of tone and rime in Cantonese sentence context was related to orthographic-to-pronunciation mapping teaching in Hong Kong and the closer link between tone and rime in Cantonese. However, these speculations still need to be examined by future studies.

4.5 Conclusion

The present study was motivated by the inconsistent findings about the integration or independence between tone and segment processing in Mandarin spoken word identification. Different from most previous studies that only concerned tone and segment processing in isolated word perception, the present study investigated the integration or independence of tone and rime processing in the context of Mandarin sentence comprehension, which is the most natural context in daily speech. The results indicated a functional dissociation of tonal and segmental information across different stages of Mandarin sentence comprehension. Specifically, rime had advantages over tone in constraining word identity, but had disadvantages in repairing. On the contrary, tone had disadvantages compared with rime in constraining word identity, but had advantages in recovering. The functional dissociation findings indicated the independence between tone and segment processing during Mandarin sentence comprehension.

References:

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. Retrieved from http://www.jstor.org/stable/2346101

Bishop, D. V. M., & Hardiman, M. J. (2010). Measurement of mismatch negativity in individuals: A study using single-trial analysis. *Psychophysiology*, *47*(4), 697–705. https://doi.org/10.1111/j.1469-8986.2009.00970.x

Brouwer, H., & Crocker, M. W. (2017). On the Proper Treatment of the N400 and P600 in Language Comprehension. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01327

Brown-Schmidt, S., & Canseco-Gonzalez, E. (2004). Who Do You Love, Your Mother or Your Horse? An Event-Related Brain Potential Analysis of Tone Processing in Mandarin Chinese. *Journal of Psycholinguistic Research*, *33*(2), 103–135. https://doi.org/10.1023/B:JOPR.0000017223.98667.10

Choi, W., & Tong, X. (2016, May 31). *Pre-attentive perceptual integration of tones and vowels*. 317–321. https://doi.org/10.21437/SpeechProsody.2016-65

Cutler, A., & Chen, H.-C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics*, *59*(2), 165–179. https://doi.org/10.3758/BF03211886

Gao, S., Hu, J., Gong, D., Chen, S., Kendrick, K. M., & Yao, D. (2012). Integration of Consonant and Pitch Processing as Revealed by the Absence of Additivity in Mismatch Negativity. *PLoS ONE*, *7*(5), e38289. https://doi.org/10.1371/journal.pone.0038289

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149–188. https://doi.org/10.1080/01690960902965951

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/BF02289823

Hu, J., Gao, S., Ma, W., & Yao, D. (2012). Dissociation of tone and vowel processing in Mandarin idioms. *Psychophysiology*, *49*(9), 1179–1190. https://doi.org/10.1111/j.1469-8986.2012.01406.x

Lee, C.-Y. (2007). Does Horse Activate Mother? Processing Lexical Tone in Form Priming. *Language and Speech*, *50*(1), 101–123. https://doi.org/10.1177/00238309070500010501

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. https://doi.org/10.1371/journal.pone.0085683

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. https://doi.org/10.1080/01690960600989600

Liu, Y., Shu, H., & Li, P. (2007). *Word naming and psycholinguistic norms: Chinese*. 7.

Luo, H., Ni, J.-T., Li, Z.-H., Li, X.-O., Zhang, D.-R., Zeng, F.-G., & Chen, L. (2006). Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proceedings of the National Academy of Sciences*, *103*(51), 19558–19563. https://doi.org/10.1073/pnas.0607065104

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. https://doi.org/10.1016/j.jml.2010.02.004

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

Paavilainen, P., Valppu, S., & Näätänen, R. (2001). The additivity of the auditory feature analysis in the human brain as indexed by the mismatch negativity: 1+1≈2 but 1+1+1<3. *Neuroscience Letters*, *301*(3), 179–182. https://doi.org/10.1016/S0304-3940(01)01635-4

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. https://doi.org/10.1162/0898929052880057

Shao, J., & Zhang, C. (2019). Talker normalization in typical Cantonese-speaking listeners and congenital amusics: Evidence from event-related potentials. *NeuroImage: Clinical*, *23*, 101814. https://doi.org/10.1016/j.nicl.2019.101814

Shuai, L., & Gong, T. (2014). Temporal relation between top-down and bottom-up processing in lexical tone perception. *Frontiers in Behavioral Neuroscience*, *8*. https://doi.org/10.3389/fnbeh.2014.00097

Singh, L., Lee, Q., & Goh, W. D. (2011). PROCESSING DEPENDENCIES BETWEEN SUPRASEGMENTAL AND SEGMENTAL INFORMATION: EFFECTS OF EMOTION AND LEXICAL TONE ON INTERFERENCE. *Hong Kong*, 4.

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. https://doi.org/10.1044/2014_JSLHR-S-13-0145

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. https://doi.org/10.1080/01690960701728261

Wiener, S., & Turnbull, R. (2016). Constraints of Tones, Vowels and Consonants on Lexical Selection in Mandarin Chinese. *Language and Speech*, *59*(1), 59–82. https://doi.org/10.1177/0023830915578000

Ye, Y., & Connine, C. M. (1999). Processing Spoken Chinese: The Role of Tone Information. *Language and Cognitive Processes*, *14*(5–6), 609–630. https://doi.org/10.1080/016909699386202

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. https://doi.org/10.1016/j.jecp.2004.05.001

Shu, H., Peng, H., & McBride-Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science*, *11*(1), 171–181. https://doi.org/10.1111/j.1467-7687.2007.00654.x

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 394–417. https://doi.org/10.1037/0278-7393.25.2.394

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. https://doi.org/10.1371/journal.pone.0142060

Gaspar, J. M., Christie, G. J., Prime, D. J., Jolicœur, P., & McDonald, J. J. (2016). Inability to suppress salient distractors predicts low visual working memory capacity. Proceedings of the National Academy of Sciences, 113(13), 3693–3698. https://doi.org/10.1073/pnas.1523471113

Zou, Y., Tsang, Y.-K., & Wu, Y. (2019). Semantic Radical Activation in Chinese Phonogram Recognition: Evidence from Event-Related Potential Recording. *Neuroscience*. https://doi.org/10.1016/j.neuroscience.2019.08.008

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. https://doi.org/10.3758/s13428-018-1038-3

Luck SJ. (2014) *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press, 2014.

CURRICULUM VITAE

Academic qualification of the thesis author, Mr. ZOU Yun:

- Received the degree of Bachelor of Science from Northeast Normal University, June 2017