Does phonological rule of tone substitution modulate mismatch negativity?

Claire H.C. Chang[a],[1], Tzu-Hui Lin[a],[1], Wen-Jui Kuo[a],[b],*

[a] *Institute of Neuroscience, National Yang-Ming University, Taipei, Taiwan*
[b] *Brain Research Center, National Yang-Ming University, Taipei, Taiwan*

ABSTRACT

This study examined whether the phonological substitution rule of tone sandhi modulates tone perception in the preattentive stage. Tone sandhi is commonly present in East Asian languages. An example from Mandarin is the Tone tone 3 sandhi rule: T3 is pronounced as T2 when followed by another T3 ($33 \rightarrow 23$). Previous mismatch negativity (MMN) studies in Mandarin have reported a smaller amplitude or longer latency in standard-deviant pair consisting of T2 and T3 (T2-T3) than in T1-T3. The most widely accepted explanation for this is that T2 and T3 have steeper pitch slopes than T1. This study tested an alternative account based on the phonological rule that the frequent substitution that occurs between T2 and T3 results in reduced MMN. In Experiment 1, we first tried to replicate the finding in Mandarin. In Experiment 2, using both unskilled and skilled speakers, we tested a sandhi tone pair of very different pitch slopes in Taiwanese. Delayed peak latency of sandhi pair was evident in both languages but only in skilled speakers. Our results did not support the shared-pitch-slope account and were instead consistent with the argument that a language-specific phonological rule could modulate preattentive tone processing.

1. Introduction

Pitch is a major information-bearing component of speech. In tone languages, pitch patterns are used to distinguish word meanings; this is referred to as lexical tone. Tone languages account for at least 40% of all languages and cluster in specific geographic regions of the world such as sub-Saharan Africa and East and Southeast Asia (Hyman, 2011, pp. 197–239; Maddieson, 2013). In this study, we adopted the mismatch negativity (MMN) paradigm and examined whether the phonological substitution rule of tone sandhi modulates tone perception in the preattentive stage. Particularly, we examined whether a standard–deviant pair consisting of tones frequently substituted for each other according to the phonological rule elicit reduced MMN.

Tone sandhi is a common feature of tone languages (Chen, 2000). An example from Mandarin is the Tone 3 sandhi rule. Mandarin has four tones, and Tone 3 (T3) is pronounced as tone 2 (T2) when it is followed by another Tone 3 ($33 \rightarrow 23$; Chao, 1948). Among the six possible tone pairs in Mandarin, T2-T3 has often been reported to be one of the most difficult pairs to discriminate for non-native speakers (Hao, 2018; Huang & Johnson, 2011; So & Best, 2014) and even more so for native speakers (Chen, Liu, & Kager, 2016, 2015; Huang & Johnson, 2011). Because non-native speakers do not know the sandhi rule, acoustic similarity is more likely the

---

reason for their difficulty. However, while acquiring a language usually leads to a better ability to discriminate acoustically similar but linguistically distinctive sounds, several behavioral studies have demonstrated that native speaker discriminated T2 and T3 slower or less accurately than non-native speakers (Chen et al., 2016, 2015; Huang & Johnson, 2011). It was suggested that this effect resulted from the acquisition of tone 3 sandhi. Huang and Johnson (2011) demonstrated that although Chinese speakers generally discriminated Chinese tones better than English speakers, for T2 and T3, Chinese speakers exhibited a longer reaction time (RT). Chen, Liu, and Kager (2015, 2016) also determined that Dutch speakers outperformed Chinese speakers at discriminating two-tone sequence consisted of T3 (33) from sequences containing T2 (tone sequence 23, 32, and 22). Therefore, acoustic similarity and the phonological rule might both increase the difficulty of discriminating T2 and T3.

Why does a phonological rule that increases perceptual ambiguity come to exist in the first place? In context, the ambiguity could be resolved with phonological, semantic, and syntactic information (Speer, Shih, & Slowiaczek, 1989, 2016), similar to the dis-ambiguation of homophones. It is possible that tone sandhi increased the ease of articulation or perception in the past but became overgeneralized and interpreted to be a categorical rule over time (Anderson, 1981; Blevins, 2006; Ohala, 1993). The pitch patterns of lexical tones in East Asian languages have undergone diachronic change and vary between dialects. Tone sandhi might have remained as a categorical phonological rule even after losing its phonetic function because of tone pattern change (Shen, 1992; Yan & Zhang, 2016; Zhang & Lai, 2010).

We wondered whether the frequent substitution between T2 and T3 modulates preattentive auditory processing and planned to test it using the MMN paradigm. Other MMN studies in Mandarin have demonstrated that the MMN elicited by the standard–deviant pair consisting of T2 and T3 was lower in amplitude and longer in peak latency than a non-sandhi tone pair (Chandrasekaran, Gandour, & Krishnan, 2007; Chandrasekaran, Krishnan, & Gandour, 2007b; Cheng et al., 2013; Hsu, Lin, Hsu, & Lee, 2014; Xi, Kager, & Gu, 2016; Li & Chen, 2015). Chandrasekaran et al. (2007; 2007a) also recruited English speakers and identified an interaction between the subject group and tone pair concerning the MMN amplitude. Only Chinese speakers exhibited weaker MMN for a sandhi tone pair than for non-sandhi tone pair, although the difference between the Chinese and English groups was not significant for either pair. One acoustic feature shared by the sandhi pair (T2-T3) but not the non-sandhi pair (T1-T3 and T1-T2) was the pitch slope. T2 and T3 both have non-flat slopes. Because the distinction between flat and non-flat pitch slopes has been reported to play a major role in dissimilarity judgment for tone language speakers than non-tone language speakers (Gandour, 1983, 1984), it is broadly accepted to be the cause of the reduced MMN (Chandrasekaran et al., 2007; Chandrasekaran et al., 2007b; Cheng et al., 2013; Hsu et al., 2014; Yu, Shafer, & Sussman, 2017).

However, although T2 and T3 both have non-flat pitch slopes, they differ in the direction. T2 is a rising tone and T3 is a low–falling (or falling–rising) tone (Fig. 2). Previous studies have reported that Mandarin speakers were more sensitive to pitch direction than English speaker (Gandour, 1983, 1984). In addition, acoustic similarity can barely explain the behavioral findings that T2 and T3 were perceived to be more similar to native than to non-native speakers (Chen et al., 2015, 2016; Huang & Johnson, 2011). No matter how similar two speech sounds are along one acoustic dimension, it is unlikely that learning a language would increase the difficulty in distinguishing phonemes/tonemes in that language. Therefore, the alternative phonological rule account is worth more consideration and examination (Li & Chen, 2015). The MMN response has been proposed to reflect the discrepancy between the deviant sound and the short-term memory trace of or the contextual expectation based on the standard sound (Naatanen, Paavilainen, Rinne, & Alho, 2007). If Mandarin T3 and T2 activate each other, their contrast may result in less discrepancy.

Another explanation for the reduced MMN comes from underspecification theory (Archangeli, 1988). According to this theory, some phonemes are not fully represented in memory. As a consequence, underspecified phonemes are often replaced or assimilated by other phonemes. In this case, when an underspecified phoneme serves as the standard, the MMN response might be smaller because of less conflict between the underlying representation of the standard and the incoming deviant sound (Cornell, Lahiri, & Eulitz, 2008; Eulitz & Lahiri, 2004); for example, the mid-vowel/ɛ/in English (Scharinger, Monahan, & Idsardi, 2016) or T3 in Mandarin Chinese (Politzer-Ahles, Schluter, Wu, & Almeida, 2016).

To better disentangle these accounts, this study included both Mandarin and Taiwanese materials. We first replicated the Mandarin findings and used similar experimental parameters for the Taiwanese experiment, so if our results were different from the Mandarin literature, changes in experimental parameters could not be the reason. Taiwanese is also called Southern Min or Minnan. It has seven tones, each of which is substituted by another tone when in a non-final position, regardless of the identity of the following
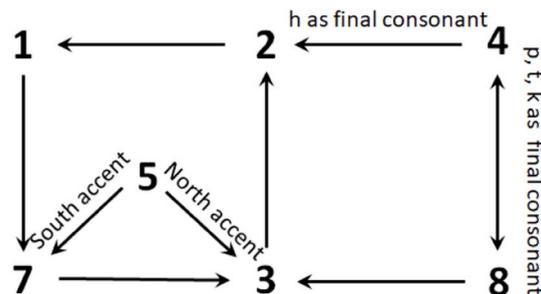

Fig. 1. Tone sandhi rules in Taiwanese. Each tone is substituted by another tone in the non-final position, regardless of the identity of the following tone. For example, T7 on the first syllable of a disyllable word is pronounced as T3 (7X→3X).
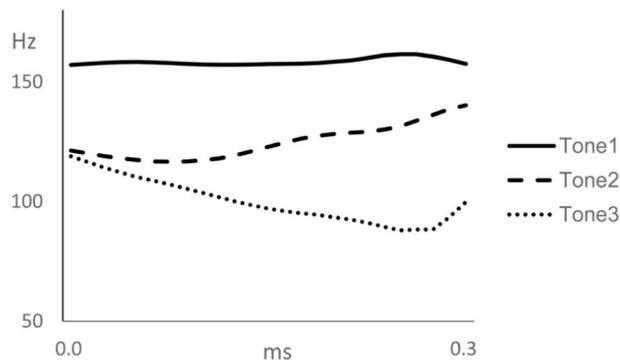
Fig. 2. Pitch contours of the three Mandarin lexical tones used as stimuli in Experiment 1. The sandhi pair consisted of T2 and T3. The control pair consisted of T1 and T3.

tone (National Languages Committee, 2011; Peng, 1997). For example, T7 on the first syllable of a disyllable word would be pronounced as T3, no matter what the following tone is (7X→3X; Fig. 1). The rich sandhi rules in Taiwanese gave us more options from which to choose materials. We selected T3 and T7 as our sandhi pair. Because T3 has a non-flat pitch slope, but T7 is flat, the reduced MMN from this pair could not be the result of shared non-flat slope (Fig. 7). Another advantage of using Taiwanese materials was that because all seven tones are replaced by other tones in the non-final position, they are all underspecified. The tone pair effect could not come from a difference in whether they are underspecified or not. Most importantly, we tested the tone sandhi effect in Taiwanese in both skilled and unskilled speakers. Differences between tone pairs evident in both skilled and unskilled groups could be explained by acoustic similarity, but the effect evident only in the skilled group must have resulted from the acquisition of Taiwanese. We expected to discover the reduced MMN for the sandhi tone pair only in the skilled group. Weaker MMN in the skilled rather than the unskilled group for the sandhi pair would strongly support the phonological rule account.

2. Experiment 1: Mandarin

2.1. Methods

2.1.1. Participants

Twelve participants with no history of neurological or psychiatric disease were recruited for the current study (six males; age range: 21–32 years; mean age: 26 years; all right-handed). They were screened at a hearing level of 30 dB at the frequencies of 500 Hz, 1 KHz, 2 KHz, and 4 KHz. All of them were native Mandarin Chinese speakers. Their written informed consent was obtained before the experiment. The current study was approved by the Institutional Review Board of National Yang-Ming University.

2.1.2. Stimuli

The combinations of a Mandarin syllable/i/(syllables are represented by international phonetic symbols in the current study) and three Mandarin lexical tones (T1, T2, and T3) were used as stimuli (/i1/"cloth,"/i2/"aunt," and/i3/"chair"; see Fig. 2 and Table 1). The sandhi pair consisted of T2 and T3. The control pair consisted of T1 and T3. The low–falling contour of the monosyllable T3 in our stimuli was slightly different from the falling–rising pattern in standard Mandarin but consistent with other studies of Mandarin spoken in Taiwan (Chang, 2010; Li, Xiong, & Wang, 2006), which might reflect the influence from Taiwanese dialect. All auditory stimuli were the recorded voice of a native male Mandarin speaker. The durations of the stimuli were adjusted to 300 ms using Audacity (http://www.audacityteam.org/), and the intensity was adjusted to 60 dB using Praat (Boersma & Weenink, 2007).

2.1.3. Design and procedure

Participants were seated in a soundproof and electrically shielded chamber. During the experiment, they passively listened to the auditory stimulus played through inserted earphones while watching a video of paper-folding art without sound or subtitles. They were required to answer questions about the video after the event-related potential (ERP) experiment; for example, "please describe the artworks in the video."

The MMN was obtained using the oddball paradigm, which included a standard sound (S) and a deviant sound (D). Four

Table 1
Pitch properties of Mandarin tone stimuli (unit: Hz).

	Onset	Offset	Average	Standard Deviation
Tone 1	156	154	158	1.93
Tone 2	123	142	126	8.34
Tone 3	123	93	101	11.29

experimental sessions were conducted: S1-D3, S3-D1, S2-D3, and S3-D2; two for the sandhi pair condition and two for the control pair condition. Each session consisted of 600 trials. The stimuli were presented using E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The interstimulus interval was 500 ms. The first nine trials were standard trials. Between any two deviant stimuli, at least two standard stimuli were presented. A standard sound was played in 80% of the 600 trials, and a deviant sound in 20% of the trials.

2.1.4. Electrophysiological recording and preprocessing

Electroencephalogram (EEG) was recorded using a 64-channel EEG system (Neuroscan SynAmps2; arranged by the 10–20 international system) in the DC mode (low-pass 100 Hz; digitized at the sampling rate of 1000 Hz) and with a reference electrode between Cz and CPz. Vertical and horizontal eye movements were measured using two pairs of bipolar channels. Another pair of bipolar electrodes was attached to the left and right mastoids. The impedances of all electrodes were kept below 10 kΩ during the experiment.

The EEG signals were segmented into epochs from −100 ms to 700 ms, relative to stimulus onset, band-pass filtered at 1–40 Hz (zero phase shifting, 12 dB/oct), and referenced to linked mastoids offline. Values between −100 ms and 0 ms for the prestimulus interval were taken as baseline. Epochs with signals exceeding ± 70 μV were excluded. To equalize the number of standard and deviant trials, only standard trials immediately followed by deviant trials were included for further analysis. The average number of accepted trials per participant per condition was 174 and the minimum was 109.

Averaged waveforms for the standard and deviant trials in each session were computed for each participant. Standard waveforms were subtracted from the deviant waveforms to obtain the difference waveforms. The difference waveform of the sandhi pair was obtained by averaging the waveforms of S2-D3 and S3-D2 sessions, and the difference waveform of the control pair was obtained by averaging the waveforms of S1-D3 and S3-D1 sessions.

2.1.5. ERP analysis

Using the FieldTrip toolbox (http://www.ru.nl/fcdonders/fieldtrip), nonparametric cluster-based analyses (Maris & Oostenveld, 2007) were performed to verify the occurrence of the MMN and to examine the difference between tone pairs. To verify the occurrence of the MMN, standard and deviant waveforms were collected. Data from each of the 801 time points and each channel were defined as one sample. For each sample, a paired t-test was performed. Temporally and spatially adjacent samples exhibiting a significant difference between the standard and deviant waveforms ($p < .05$) were grouped into clusters. A cluster-level t-statistic was defined as the sum of t-values within a cluster. To verify that the observed cluster did not occur just by chance, the aforementioned procedure was repeated after shuffling the condition labels within every participant and the t-statistic of the biggest cluster was extracted. After 1000 repetitions, a distribution of the t-statistic was obtained. If the t-statistic of our cluster exceeded 95% of the distribution obtained with shuffled condition labels, the cluster was significant at a threshold of Monte Carlo $p < .05$. A similar analysis was conducted to examine the difference between tone pairs. We only reported clusters centered around 100–300 ms after stimulus onset at the frontal channels.

For the peak latency and amplitude analyses, we divided the frontal channels into left (F7, F5, F3, FT7, FC5, and FC3), middle (F1, Fz, F2, FC1, FCz, and FC2), and right (F4, F6, F8, FC4, FC6, and FT8) sites. Averaged difference waveforms from the three sites were computed for each condition and each participant. MMN amplitudes were measured by finding the negative peak within the 100–300 ms time window and averaging the voltage within a 50-ms time window centered on the peak. Repeated-measure ANOVAs with the channel site (left, middle, and right) and tone pair as within-subject factors were performed using software R (R Development Core Team, 2008) and its ez package (Lawrence, 2013). Post hoc pairwise comparisons were performed using paired t-tests with the p value corrected by the Bonferroni method.

2.2. Results

2.2.1. Cluster analysis

Cluster-based analysis verified the occurrence of the MMN in the 103–256 ms time window after stimulus onset (Fig. 3).
The cluster-based analysis also indicated a significant tone pair effect on the MMN in the 78–196 ms time window (Fig. 4).

2.2.2. Peak latency and amplitude analyses

Consistent with the cluster analysis, ANOVA of the MMN peak latency indicated a main effect of tone pair (Table 2 and Fig. 5, upper panel; $F(1,11) = 28.44, p < .01$), reflecting that the MMN of the control tone pair peaked earlier than that of sandhi pair. The paired t-test indicated that the tone pair effect was significant for all three sites with Bonferroni correction (left: $t(1,11) = 6.02$, $p < .01$; middle: $t(1,11) = 3.39, p < .01$; right: $t(1,11) = -5.26, p < .01$).

ANOVA of the peak amplitude indicated a main effect from the channel site (Table 2 and Fig. 5, lower panel; $F(2,22) = 0.08$, $p < .01$), reflecting that the MMN at the middle site was stronger than that at the left and right sites (middle vs. left: $(1,23) = -3.78$, $p < .01$; middle vs. right: $t(1,23) = -4.69, p < .01$; right vs. left: $t(1,23) = -0.79, p > .05$).

3. Experiment 2: Taiwanese

3.1. Methods

If not otherwise stated, the methods used in Experiment 2 were the same as those described for Experiment 1.
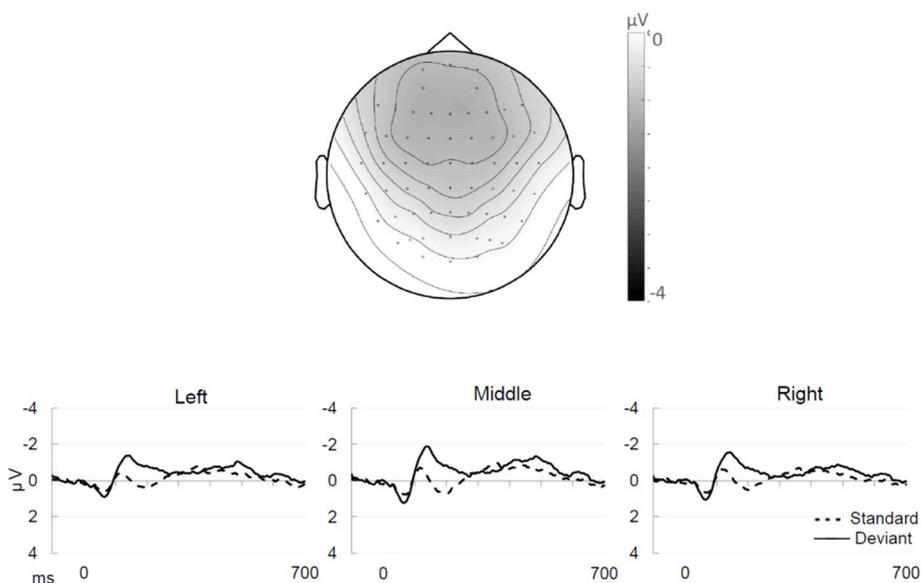
MMN response

Fig. 3. MMN response by native speakers to Mandarin tone change. Upper: Topology of the MMN response identified by cluster analysis 103–256 ms after stimulus onset. Lower: Grand average waveforms elicited by the deviant and standard stimuli in the left, middle, and right frontal sites.
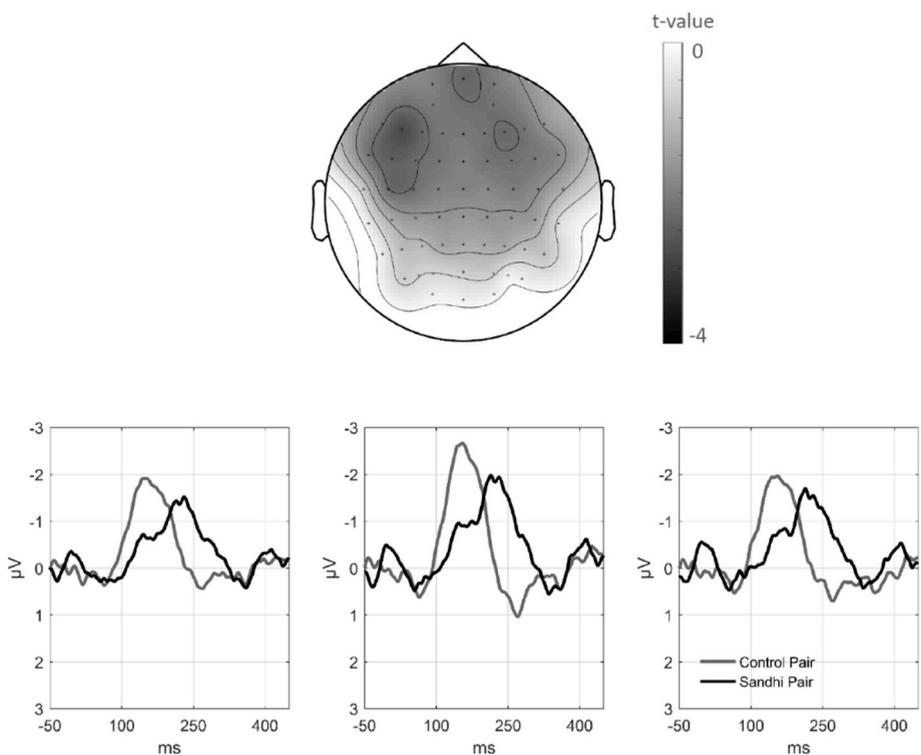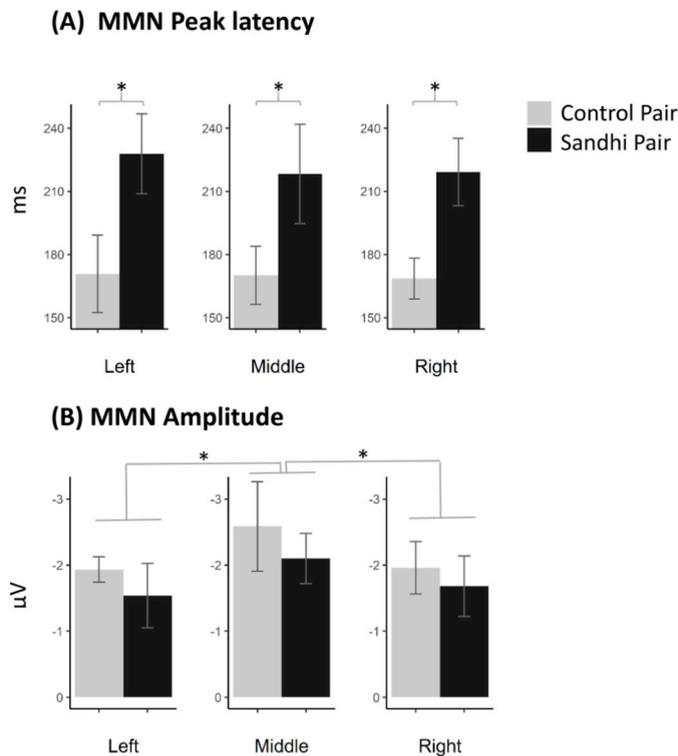
Tone pair effect

Fig. 4. Tone pair effect on the MMN. Upper: Topology of the tone pair effect identified by the cluster analysis 78–196 ms after stimulus onset. Lower: MMN waveforms of the control and sandhi tone pairs at three frontal sites.

Table 2

Tone pair effect on MMN peak latency and amplitude at three frontal sites (SD = standard deviation).

Amplitude (μV)	Left	Middle	Right
Sandhi Pair	−1.54 (SD = 0.77)	−2.10 (SD = 0.60)	−1.68 (SD = 0.72)
Control Pair	−1.93 (SD = 0.30)	−2.59 (SD = 1.07)	−1.96 (SD = 0.62)
Cohen's D	0.47	0.35	0.24
Peak latency (ms)	**Left**	**Middle**	**Right**
Sandhi Pair	228 (SD = 30)	218 (SD = 37)	219 (SD = 25)
Control Pair	171 (SD = 29)	170 (SD = 22)	169 (SD = 15)
Cohen's D	1.74	0.98	1.52

Fig. 5. Tone pair effect on MMN peak latency and amplitude at the frontal sites. Error bars represent 95% confidence interval (CI) across participants. Asterisks indicate a significant ($p < .01$) tone pair effect with Bonferroni correction.

3.1.1. Participants

Thirty participants were recruited (nine males; age range 18–32 years; mean age: 24.43 years; all right-handed). We recruited our participants in Taiwan, which is a multilingual and multiethnic society. Although Taiwanese is one of the major languages spoken in Taiwan, Mandarin is still the dominant one. Therefore, all of the participants were native Mandarin Chinese speakers. Their skills in Taiwanese were measured with a pretest.

Two participants were excluded from the statistical analysis; one with zero trials remained under one condition after EEG pre-processing and the other exhibited aberrant pretest performance (achieving 79% accuracy for sandhi application but 6% accuracy for word meaning explanation; please see below).

3.1.2. Pretest

We evaluated the participants' Taiwanese skill before the ERP experiment. The participants were asked to listen to a Taiwanese syllable and repeat the syllable twice to see whether they applied the Taiwanese sandhi rule. Then, they were asked to provide the meaning of the disyllable word they just pronounced.

Fifty-three Taiwanese syllables were used in the pretest, including 13 syllables with T1, 12 syllables with T2, 12 syllables with T3, 12 syllables with T7, and 4 syllables with T5. All the syllables and their repeated forms have corresponding words. Sound recordings of these syllables were recorded by a male native Taiwanese speaker. Every participant went through the 53 testing syllables in a random order. Each trial started with a fixation of 1000 ms, followed by the sound of 300 ms. The participants had 3000 ms to repeat
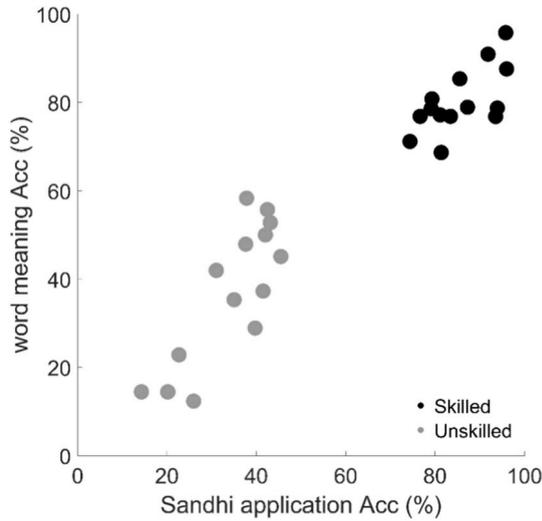
Fig. 6. Performance of the skilled and unskilled groups on the pretest.

the syllable and 5000 ms to explain the meaning of the disyllable word. Their answers were recorded.

The results indicated a strong correlation between sandhi application and word meaning accuracies across participants (r = 0.94, *p* < .001; Fig. 6). The participants formed two clear clusters, especially with respect to sandhi application accuracy. The two clusters can be separated along either dimension at 60% accuracy. We, therefore, divided our participants into skilled and unskilled groups. The average word meaning accuracy was 80% (SD: 7%) for the skilled group and 37% (SD: 16%) for the unskilled group. The average sandhi rule application accuracy was 86% (SD: 7%) for the skilled group and 34% (SD: 10%) for the unskilled group.

3.1.3. Stimuli

The combinations of two Taiwanese syllables (/*iau*/and/*uã*/) and three tones (T1, T3, and T7) were used as stimuli (/*iau2*/"still,"/*iau3*/"only need,"/*iau7*/"shining,"/*uã2*/"bowl,"/*uã3*/"evening," and/*uã7*/"drought"; Fig. 7 and Table 3). The sandhi pair consisted of T3 and T7. The control pair consisted of T2 and T7. In Taiwanese, T7 is pronounced as T3 when followed by another tone, regardless of the identity of the following tone (National Languages Committee, 2011; Peng, 1997). We chose/iau/and/uã/because, first, their combinations with T2, T3, and T7 are phonotactically valid in Taiwanese. Second, they start with vowels. Tone is clearer in vowels than in consonants. Third, their combinations with T2, T3, and T7 do not form a word when repeated (an English example: "lily"); thus, using these syllables minimized the top-down influence from word-level processing in this experiment.

3.1.4. Design and procedure

Each participant went through eight experimental sessions (S3-D7/S7-D3/S2-D7/S7-D2 x uã/iau). Each session consisted of 300 trials. The MMN scores of the control pair were obtained by averaging the difference waveforms of S2D7 and S7D2 sessions, and the MMN scores of the sandhi pair were obtained by averaging the difference waveforms of S3-D7 and S7-D3 sessions.

3.1.5. ERP analysis

The average number of accepted trials per participant per condition was 92, and the minimum was 49. The cluster-based analysis was performed for the skilled and unskilled groups separately to verify the occurrence of the MMN. The cluster-based analysis was
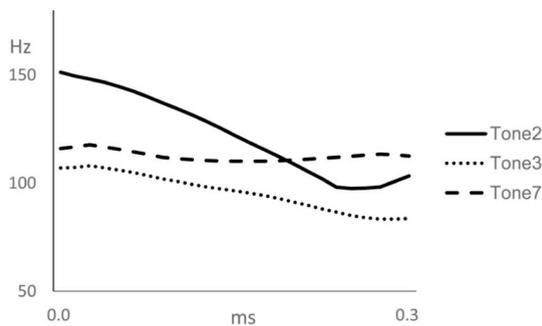

Fig. 7. Pitch contours of the three Taiwanese lexical tones used as stimuli in Experiment 2. The sandhi pair consisted of T3 and T7. The control pair consisted of T2 and T7.

Table 3

Pitch properties of Taiwanese tone stimuli (unit: Hz).

	Onset	Offset	Average	Standard Deviation
Tone 2	157	103	123	20.26
Tone 3	106	85	96	9.00
Tone 7	114	110	113	2.62

also performed to test the tone pair effect and the interaction between the tone pair and syllable for each group. For the interaction, instead of a paired t-test, Hotelling's T-squared test was applied to each sample, as suggested in http://www.fieldtriptoolbox.org/faq/how_can_i_test_an_interaction_effect_using_cluster-based_permutation_tests. The cluster level t-statistic was defined as the sum of t-squared within a cluster.

3.2. Results

3.2.1. Cluster analysis

Cluster-based analysis indicated a typical MMN response at the frontal sites in the 124–286 ms time window for the unskilled group (Fig. 8) and in the 115–299 ms time window for the skilled group (Fig. 9).

The tone pair effect was not significant for either group in cluster-based analyses. However, a significant tone pair X syllable interaction was observed in the skilled group in the 126–227-ms time window. Further examination indicated that the tone pair effect was only significant for the syllable/uã/(130–206 ms, Fig. 10). The syllable/uã/is only a legal syllable in Taiwanese, but/iau/is legal in both Mandarin and Taiwanese. Because the participants in the skilled group were native speakers of both Mandarin and Taiwanese, the syllable/uã/might provide a better context for Taiwanese sandhi rule.

3.2.2. Peak latency and amplitude analyses

Because cluster-based analysis indicated a tone pair effect only for the syllable/uã/, we focused on/uã/in the latency and amplitude analyses (Fig. 11). Mixed ANOVA with skill as the between-subject factor and tone pair and channel site as within-subject factors was performed. For peak latency, we identified a significant three-way interaction between skill, tone pair, and channel site ($F_{(2,52)} = 6.72$, $p = .003$; Table 4 and Fig. 12, upper panel). The two-way interaction between skill and tone pair was not significant when tested for each site separately (left site: $F_{(1,26)} = 3.53$, $p = .07$; middle: $F_{(1,26)} = 1.02$, $p = .32$; right: $F_{(1,26)} = 0.004$, $p = .94$). Post hoc paired t-tests of the tone pair effect were conducted with Bonferroni correction. For the unskilled group, the tone
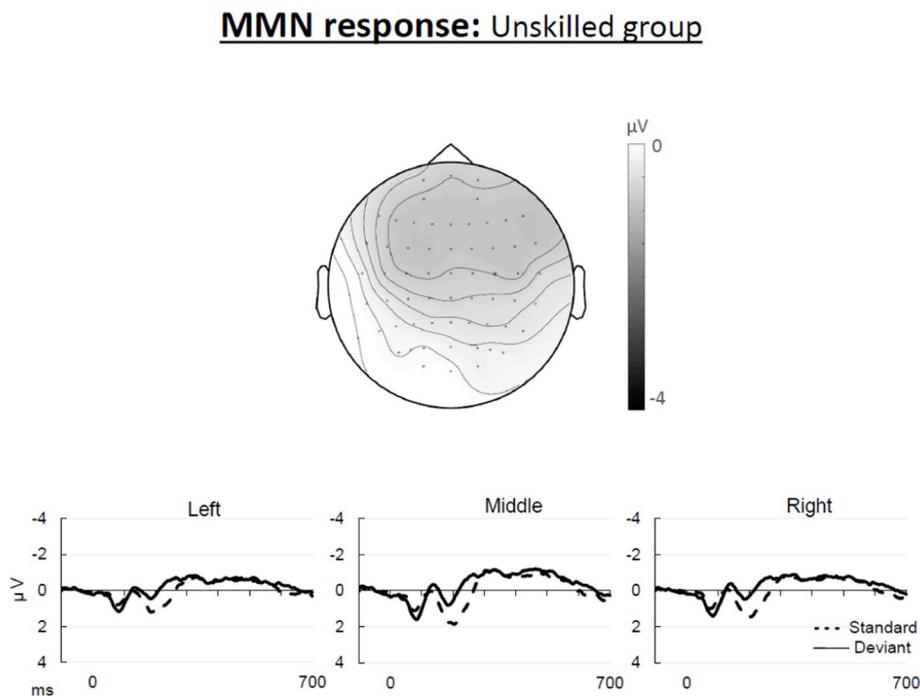

Fig. 8. MMN response of unskilled speakers to Taiwanese tone change. Upper: Topology of the MMN response identified by cluster analysis 124–286 ms after stimulus onset. Lower: Grand average waveforms elicited by the deviant and standard stimuli in the left, middle, and right frontal sites.

MMN response: Skilled Group
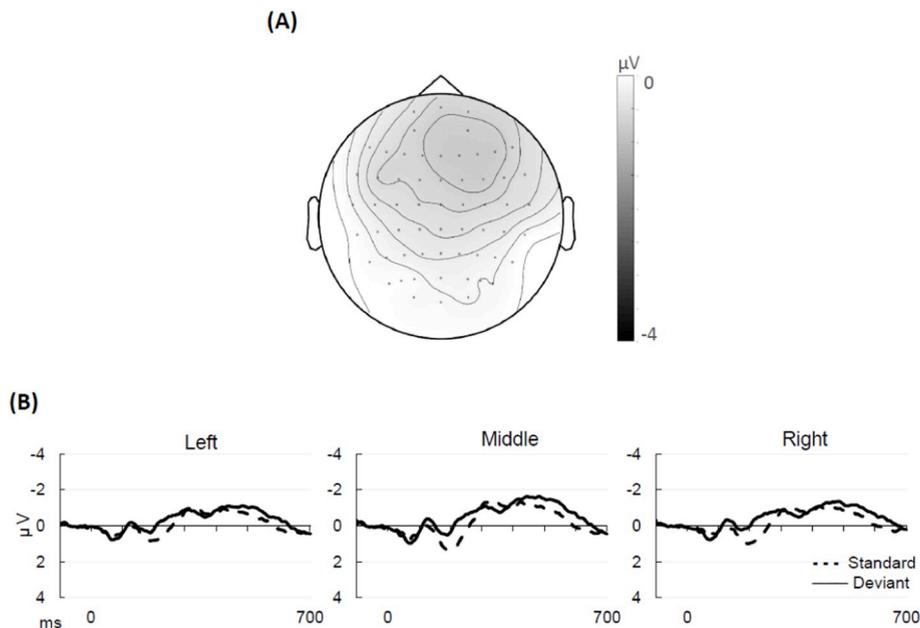
(A)

(B)

Fig. 9. MMN response of skilled speakers to Taiwanese tone change. Upper: Topology of the MMN response identified by cluster analysis 115–299 ms after stimulus onset. Lower: Grand average waveforms elicited by the deviant and standard stimuli in the left, middle, and right frontal sites.

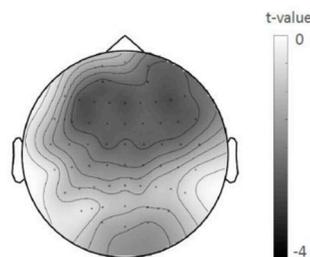**Tone pair effect:** Skilled group

Fig. 10. Topology of the tone pair effect on the MMN of the skilled group for the syllable/uã/, as identified by the cluster analysis 130–206 ms after stimulus onset.

pair effect was not significant under any conditions. For the skilled group, the tone pair effect was significant at the left (t (1,13) = 6.87, $p < .01$) and middle sites (t(1,13) = 5.99, $p < .01$) and insignificant at the right site (t(1,13) = 3.61, $p > .05$). However, post hoc two-sample t-tests of the group effect were not significant under any conditions.

For peak amplitude, we only discerned a channel site effect but no significant effect with respect to tone pair, skill, or their interaction. A post hoc paired t-test indicated that MMN amplitude was largest at the middle site and smallest at the left site (Table 4 and Fig. 12, lower panel; middle vs. right t(1,55) = −5.40, $p < .01$; right vs. left: t(1,55) = −2.85, $p < .05$; middle vs. left: t (1,55) = −7.50, $p < .01$).

4. Discussion

This study aimed to examine whether a language-specific phonological rule modulates the preattentive processing of a lexical tone. Specifically, we predicted that the sandhi rule that involves tone substitution would lead to reduced or delayed MMN for the sandhi tone pair in native speakers. Experiment 1 demonstrated that the peak latency of the MMN to the sandhi tone pair was longer than that to the control pair (Figs. 4 and 5), consistent with other studies of Mandarin (Chandrasekaran et al., 2007; Chandrasekaran, Krishnan, et al., 2007b; Cheng et al., 2013; Hsu et al., 2014; Li & Chen, 2015). Three possible explanations exist for the reduced MMN
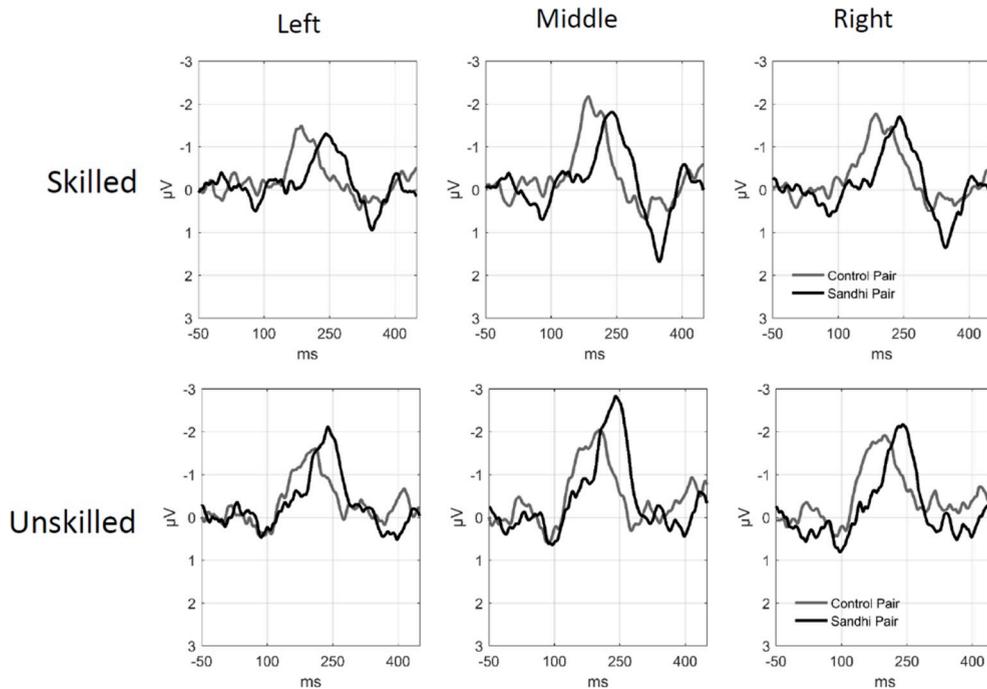
Fig. 11. MMN waveforms of the control and sandhi tone pairs at three frontal sites for the syllable/uã/.

Table 4
Tone pair effect on MMN peak latency and amplitude at three frontal sites for the syllable/uã/.

Peak Latency (ms)		Left	Middle	Right
Skilled	Sandhi Pair	240 (SD = 31)	231 (SD = 37)	225 (SD = 32)
	Control Pair	185 (SD = 26)	184 (SD = 23)	194 (SD = 24)
	Cohen's D	1.84	1.60	0.96
Uskilled	Sandhi Pair	234 (SD = 20)	234 (SD = 29)	237 (SD = 23)
	Control Pair	204 (SD = 30)	203 (SD = 34)	205 (SD = 40)
	Cohen's D	0.73	0.57	0.61
Amplitude (μV)		**Left**	**Middle**	**Right**
Skilled	Sandhi Pair	−1.55 (SD = 0.57)	−2.15 (SD = 0.83)	−1.87 (SD = 0.69)
	Control Pair	−1.40 (SD = 0.65)	−1.95 (SD = 0.54)	−1.65 (SD = 0.64)
	Cohen's D	0.16	0.17	0.20
Unskilled	Sandhi Pair	−2.01 (SD = 0.81)	−2.78 (SD = 0.86)	−2.56 (SD = 0.46)
	Control Pair	−1.99 (SD = 0.68)	−2.67 (SD = 0.75)	−2.02 (SD = 0.70)
	Cohen's D	0.03	0.13	0.79

with respect to the contrast between Mandarin T2 and T3: the phonological rule account, the underspecification account, and the acoustic similarity account (e.g., pitch slope).

To disentangle these accounts, we used Taiwanese materials in Experiment 2 and compared skilled and unskilled groups of speakers. We discovered that the sandhi pair yielded a delayed MMN peak latency in the skilled group but not the unskilled group (Figs. 11 and 12). These results did not support the underspecification account. Because all the Taiwanese tones are underspecified (i.e., they are all substituted by other tones in the non-final position), the tone pair effect could not reflect the difference in whether they were underspecified or not. Further, our Taiwanese sandhi pair consisted of T3, a contour tone with a non-flat pitch slope, and T7, a level tone with a flat pitch slope, and the reduced MMN could not result from a shared non-flat pitch slope. However, we could not completely rule out the possibility that skilled speakers were just more sensitive to other acoustic features shared by the sandhi tone pair. For example, the starting pitch height of Taiwanese T7 is closer to T3 than T2, and the starting pitch of Mandarin T3 is closer to T2 than T1. Although behavioral studies have demonstrated that non-tone language speakers are more sensitive to average pitch height than tone language speakers (Gandour, 1983; Gandour & Harshman, 1978; Jongman, Qin, Zhang, & Sereno, 2017), our data were not sufficient to discard pitch height as a potential explanation for the delayed MMN. Our results were consistent with the phonological rule account. In particular, the reduced MMN resulted from the frequent tone substitution required by the sandhi rule. Our results also suggested that this effect is language specific. The skilled and unskilled groups were both native tone language
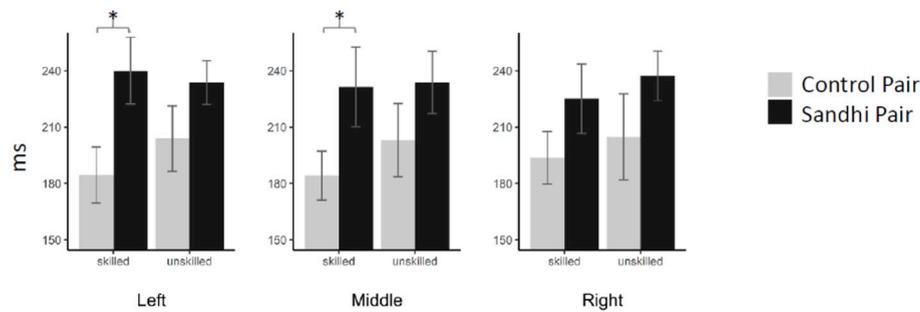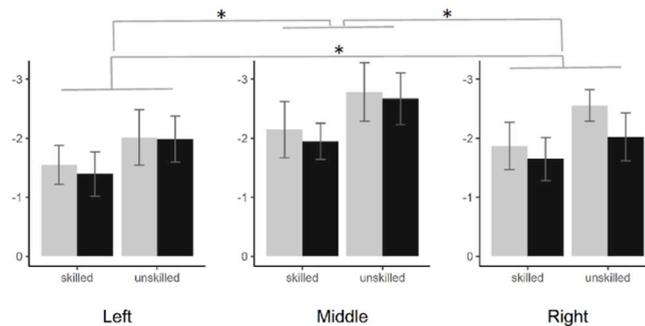
(A) MMN Peak Latency

(B) MMN Peak Amplitude

Fig. 12. Tone pair effect on MMN peak latency and amplitude at three frontal sites for the syllable/uã/. Error bars represent 95% CI across participants. Asterisks indicate a significant ($p < .01$) tone pair effect with Bonferroni correction.

speakers (Mandarin and Taiwanese); thus, if the difference between the sandhi and control pairs had resulted from varying degrees of experience with tone languages, both groups should have exhibited the effect. However, this effect was only evident in skilled Taiwanese speakers; thus, experience in any tone languages is not enough. Only experience in a tone language having the tested sandhi rule—Taiwanese—can induce this effect.

The tone pair effect in Taiwanese was evident for the syllable/uã/but not/iau/. The reason might have been that our skilled Taiwanese speakers were also native Chinese speakers and/iau/is a legal syllable in both Mandarin and Taiwanese. Therefore,/uã/ provided a better context for Taiwanese tone sandhi than/iau/, meaning that the effect of language experience on tone perception might depend on the context. Huang and Johnson (2011) reported that Chinese speakers discriminated natural speech tones better than English speakers, but English speakers discriminated Mandarin tones borne by sine wave better than Chinese speakers. Although Chinese speakers have more experience with tones, this advantage depends on the context (speech sound vs. sine wave). Chen et al. (2016) reported that Dutch speakers outperformed Chinese speakers at discriminating the sandhi tone pair in disyllabic stimuli, which was a better context for Mandarin Tone 3 sandhi ($33 \rightarrow 23$) than monosyllabic stimuli. In the future, manipulating the linguistic context systematically (e.g., pseudowords vs. nonwords vs. nonspeech) is a potential method to further disentangle the effects of a phonological rule and acoustic similarity. The acoustic similarity effect supposedly depends less on the linguistic context.

In addition to tone, phonological rules of phoneme change (i.e., place assimilation; e.g.,/d/to/b/in "bad boy") have been demonstrated to modulate the MMN effect as well (Mitterer & Blomert, 2003; Mitterer, Csépe, Honbolygo, & Blomert, 2006; Sun et al., 2015; Tavabi, Elling, Dobel, Pantev, & Zwitserlood, 2009). However, place assimilation is usually optional and not language specific, reflecting the universal demand on the ease of articulation. By contrast, little evidence exists concerning the premise that the tone sandhi rules tested in this study can increase the ease of articulation (Myers & Tsay, 2008; Shih, 2008; Xu, 2004; Jongman et al., 2017.; Zhang & Lai, 2010). Because many tone sandhi rules are obligatory and language specific (Chen, 2000), they can serve as probes for examining the top-down influence of a high-level phonological rule on preattentive auditory processing.

The nature of tone sandhi depends on the exact rule in question and could vary between languages. It has been argued that Mandarin and Taiwanese tone sandhi are different, and Mandarin is more generalized and productive (Tsay & Myers, 1996). The key direct evidence for this claim came from the observation that Mandarin speakers applied the Mandarin sandhi rule to pseudo-words (Zhang & Lai, 2010; Zhang & Peng, 2013) more often than Taiwanese speakers (Hsieh, 1970, 1975, 1976; Wang, 1993; Jongman et al., 2017; Zhang & Lai, 2007). Whether such a difference affects, the preattentive stage of auditory processing could be tested by using pseudo-words in future MMN studies. If the sandhi effect was evident even in pseudo-word in Mandarin but not in Taiwanese, the contentions that the Mandarin sandhi rule is more generalized and such a difference affects early auditory processing would be supported.

In summary, we discovered delayed MMN to a sandhi tone pair (i.e., tone pairs often substituted for each other according to a phonological rule) both in Mandarin and Taiwanese. Further, such an effect was only evident in skilled speakers. Our results were consistent with the argument that a language-specific phonological rule could modulate preattentive tone processing. Our results did not support the underspecification and shared non-flat slope accounts because the Taiwanese stimuli did not differ in the degree of underspecification and the two tones in the sandhi pair had very different pitch slopes. However, further study is needed to examine whether skilled speakers are just more sensitive to other acoustic features shared by the sandhi pair.

Acknowledgments

References

Anderson, S. R. (1981). *Why Phonology Isn't "Natural.", 12*(4), 493–539.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology, 5*(02), 183–207. https://doi.org/10.1017/S0952675700002268.

Blevins, J. (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics, 32*(2), 117–166. https://doi.org/10.1515/TL.2006.009.

Boersma, P., & Weenink, D. (2007). *Praat: doing phonetics by computer.* [Computer program] Http://Www.Praat.Org/Version 4.5.

Chandrasekaran, B., Gandour, J. T., & Krishnan, A. (2007). Neuroplasticity in the processing of pitch dimensions: a multidimensional scaling analysis of the mismatch negativity. *Restorative Neurology and Neuroscience, 25*(3–4), 195–210. https://doi.org/10.1016/j.ygyno.2014.12.035.Pharmacologic.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007a). Experience-dependent neural plasticity is sensitive to shape of pitch contours. *NeuroReport, 18*(18), 1963–1967. https://doi.org/10.1097/WNR.0b013e3282f213c5.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007b). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research, 1128*(1), 148–156. https://doi.org/10.1016/j.brainres.2006.10.064.

Chang, C. Y. (2010). *Dialect Differences in the Production and Perception of Mandarin Chinese Tones*.

Chao, Y. R. (1948). *Mandarin Primer.* Cambridge (UK): Harvard University Press. https://doi.org/10.4159/harvard.9780674732889.

Chen, M. Y. (2000). *Tone Sandhi patterns across Chinese dialects.* Cambridge University Press.

Cheng, Y.-Y., Wu, H.-C., Tzeng, Y.-L., Yang, M.-T., Zhao, L.-L., & Lee, C.-Y. (2013). The development of mismatch responses to Mandarin lexical tones in early infancy. *Developmental Neuropsychology, 38*(5), 281–300. https://doi.org/10.1080/87565641.2013.799672.

Chen, A., Liu, L., & Kager, R. (2015). Cross-linguistic perception of Mandarin tone sandhi. *Language Sciences, 48*, 62–69. https://doi.org/10.1016/j.langsci.2014.12.002.

Chen, A., Liu, L., & Kager, R. (2016). Cross-domain correlation in pitch perception, the influence of native language. *Language, Cognition and Neuroscience, 31*(6), 751–760. https://doi.org/10.1080/23273798.2016.1156715.

Cornell, S. A., Lahiri, A., & Eulitz, C. (2008). Behavioral and brain sciences, volume 31 - issue 05. *Behavioral and Brain Sciences, 31*(05), 79–89. https://doi.org/10.1017/S0140525X08004998.

Eulitz, C., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience, 16*, 577–583. https://doi.org/10.1162/089892904323057308.

Gandour, J. T. (1983). *Tone perception in far eastern-languages. Journal of Phonetics, 11*(2), 149–175.

Gandour, J. T. (1984). *Tone dissimilarity judgments by Chinese Listeners. Journal of Chinese Linguistics, 12*(2), 235–261.

Gandour, J. T., & Harshman, R. A. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech, 21*(1), 1–33. https://doi.org/10.1177/002383097802100101.

Hao, Y.-C. (2018). Second Language Perception of Mandarin Vowels and Tones. *Language and Speech, 61*(1), 135–152. https://doi.org/10.1177/0023830917717759.

Hsieh, H.-I. (1970). The psychological reality of tone sandhi rules in Taiwanese. *Papers From the 6Th Annual Regional Meeting of the Chicago Linguistic Society* (pp. 489–503). .

Hsieh, H.-I. (1975). How generative is phonology. *The Transformational-Generative Paradigm and Modern Linguistic Theory, 109–144*.

Hsieh, H.-I. (1976). On the unreality of some phonological rules. *Lingua, 38*(1), 1–19. https://doi.org/10.1016/0024-3841(76)90038-3.

Hsu, C. H., Lin, S. K., Hsu, Y. Y., & Lee, C. Y. (2014). The neural generators of the mismatch responses to Mandarin lexical tones: An MEG study. *Brain Research, 1582*, 154–166. https://doi.org/10.1016/j.brainres.2014.07.023.

Huang, T., & Johnson, K. (2011). Language specificity in speech perception: Perception of mandarin tones by native and nonnative listeners. *Phonetica, 67*(4), 243–267. https://doi.org/10.1159/000327392.

Hyman, L. L. M. (2011). *Tone: Is it Different? The Handbook of Phonological Theory* (2nd ed.). . (1997) https://doi.org/10.1002/9781444343069.ch7.

Jongman, A., Qin, Z., Zhang, J., & Sereno, J. A. (2017). Just noticeable differences for pitch direction, height, and slope for Mandarin and English listeners. *Journal of the Acoustical Society of America, 142*(2), EL163–EL169. https://doi.org/10.1121/1.4995526.

Lawrence, M. A. (2013). *Easy analysis and visualization of factorial experiments*.

Li, X., & Chen, Y. (2015). Representation and processing of lexical tone and tonal variants: Evidence from the mismatch negativity. *PLoS One, 10*(12), 1–24. https://doi.org/10.1371/journal.pone.0143097.

Li, A., Xiong, Z., & Wang, X. (2006). Contrastive study on tonal patterns between accented and standard Chinese. *Proceedings of the 5th International Symposium on Chinese Spoken Language. Vol. 2006. Proceedings of the 5th International Symposium on Chinese Spoken Language* (pp. 157–168).

Maddieson, I. (2013). Tone. In M. S. Dryer, & M. Haspelmath (Eds.). *The World Atlas of Language Structures.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & Psychophysics, 65*(6), 956–969. https://doi.org/10.3758/BF03194826.

Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The Recognition of Phonologically Assimilated Words Does Not Depend on Specific Language Experience. *Cognitive Science, 30*(3), 451–479. https://doi.org/10.1207/s15516709cog0000_57.

Myers, J., & Tsay, J. (2008). Neutralization in Taiwan Southern Min Tone sandhi. *Interfaces in Chinese Phonology: Festschrift in Honor of Matthew Y. Chen on His 70th Birthday, (March 2002)* (pp. 47–78). . https://doi.org/10.1002/hbm.21191.

Naatanen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026.

National Languages Committee (2011). *Online practical dictionary of Taiwanese.* http://twblg.dict.edu.tw/holodict_new/compile1_3_9_3.jsp.

Ohala, J. J. (1993). Coarticulation and phonology. *Language and Speech, 36*https://doi.org/10.1177/002383099303600303.

Peng, S. (1997). Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics, 25*(3), 371–400. https://doi.org/https://doi.org/10.1006/jpho.1997.0047.

Politzer-Ahles, S., Schluter, K., Wu, K., & Almeida, D. (2016). Asymmetries in the perception of Mandarin tones: Evidence from mismatch negativity. *Journal of*

Experimental Psychology: Human Perception and Performance, 42(10), 1547–1570. https://doi.org/10.1037/xhp0000242.

R Development Core Team, R. (2008). *Computational Many-Particle Physics.* Vienna, Austria: R Foundation for Statistical Computing. R Foundation for Statistical Computing https://doi.org/10.1007/978-3-540-74686-7.

Scharinger, M., Monahan, P. J., & Idsardi, W. J. (2016). Linguistic category structure influences early auditory processing: Converging evidence from mismatch responses and cortical oscillations. *NeuroImage, 128*, 293–301. https://doi.org/10.1016/j.neuroimage.2016.01.003.

Shen, X. S. (1992). On tone sandhi and tonal coarticulation. *Acta Linguistica Hafniensia, 25*(1), 83–94. https://doi.org/10.1080/03740463.1992.10412279.

Shih, C. (2008). Linking phonology and phonetics: An implementation model of tones. In H. H. Yuchau Hsiao Lian-Hee Wee, & D.-an Ho (Vol. Eds.), *Interfaces in Chinese phonology: Festschrift in honor of Matthew Y. Chen on his 70th birthday: Vol. 0623805,* (pp. 99–120). Taipei: Academia Sinica 漢語音韻介面：陳淵泉先生七秩壽慶論文集.

So, C. K., & Best, C. T. (2014). Phonetic influences on english and french listeners' assimilation of mandarin tones to native prosodic categories. *Studies in Second Language Acquisition, 36*(2), 195–221. https://doi.org/10.1017/S0272263114000047.

Speer, S. R., Shih, C.-L. L., & Slowiaczek, M. L. (1989). Prosodic structure in language understanding: Evidence from tone sandhi in mandarin. *Language and Speech, 32*(4), 337–354. https://doi.org/10.1177/002383098903200403.

Speer, S. R., Shih, C.-L., & Slowiaczek, M. L. (2016). *Prosodic Structure in Language Understanding: Evidence from Tone Sandhi in Mandarin.* https://doi.org/10.1177/002383098903200403.

Sun, Y., Giavazzi, M., Adda-Decker, M., Barbosa, L. S., Kouider, S., Bachoud-Lévi, A. C., ... Peperkamp, S. (2015). Complex linguistic rules modulate early auditory brain responses. *Brain and Language, 149*, 55–65. (2009) https://doi.org/10.1016/j.bandl.2015.06.009.

Tavabi, K., Elling, L., Dobel, C., Pantev, C., & Zwitserlood, P. (2009). Effects of place of articulation changes on auditory neural activity: A magnetoencephalography study. *PLoS One, 4*(2)https://doi.org/10.1371/journal.pone.0004452.

Tsay, J., & Myers, J. (1996). Taiwanese Tone Sandhi as Allomorph Selection. *Proceedings of Annual Meeting of the Berkeley Linguistics Society.*

Wang, X. S. (1993). Taiyu Biandiao De Xinli Texing (on the Psychological Status of Taiwanese Tone Sandhi). *Tsinghua Xuebao, 23*(23), 175–192.

Xi, L., Kager, R., & Gu, W. (2016). Surface vs. Underlying Listening Strategies for Cross-Language Listeners in the Perception of Sandhied Tones in the Nanjing Dialect. *Proceedings of the 5th International Symposium on Tonal Aspects of Languages (TAL 2016), (May)* (pp. 33–37). .

Xu, Y. (2004). Understanding tone from the perspective of production and perception. *Language and Linguistics, 5*(4), 757–797.

Yan, H., & Zhang, J. (2016). Pattern substitution in Wuxi tone sandhi and its implication for phonological learning. *International Journal of Chinese Linguistics, 3*(1), 1–45. https://doi.org/10.1075/ijchl.3.1.01yan.

Yu, Y. H., Shafer, V. L., & Sussman, E. S. (2017). Neurophysiological and Behavioral Responses of Mandarin Lexical Tone Processing. *Frontiers in Neuroscience, 11*, 95. https://doi.org/10.3389/fnins.2017.00095.

Zhang, J., & Lai, Y. (2007). Two aspects of productivity in Taiwanese Double Reduplication. *Kansas Working Papers in Linguistics, 29*, 33–47.

Zhang, J., & Lai, Y. (2010). *Testing the role of phonetic knowledge in Mandarin tone sandhi. Phonology*, Vol. 27University of Kansas. Linguistics Graduate Student Associationhttps://doi.org/10.1017/S0952675710000060.

Zhang, C., & Peng, G. (2013). Productivity of Mandarin Third Tone sandhi: A wug test. *Eastward Flows the Great River: Festschrift in Honor of Prof. William S-Y. Wang on His 80th Birthday, (August)* (pp. 256–282). .