Journal of Neurolinguistics

Investigating MMN Responses to Lexical Tone Contrasts in Monolingual and Bilingual Speakers of Tonal Languages

--Manuscript Draft--

Manuscript Number:	JNL-D-23-00038
Article Type:	Research Paper
Keywords:	MMN; lexical tones; tonal language; bilinguals; Mandarin Chinese; ERPs
Corresponding Author:	Chun-Hsien Hsu National Central University Taoyuan City, TAIWAN
First Author:	Chun-Hsien Hsu
Order of Authors:	Chun-Hsien Hsu
	Wen-Chun Huang
	Tong-Hou Cheong
Abstract:	The present study aimed to evaluate whether tonal language speakers' language backgrounds influence perceptual processes using an auditory oddball paradigm. Measurements of mismatch negativity (MMN) activity to tonal contrasts found that Hakka-Mandarin bilinguals showed similar MMN activity to tonal contrasts regardless of language type. Conversely, Mandarin monolingual speakers' data showed stronger MMN activity to Mandarin syllables than to Hakka syllables. The present study demonstrates that monolingual speakers might store a bank of phonological exemplars for perceiving the tonal information of native syllables, whereas bilinguals have a generic representation of pitch heights that allows them to process the tonal information of different tonal languages.
Suggested Reviewers:	Yuchun Chen, PhD Associate Professor, Fu Jen Catholic University 128162@mail.fju.edu.tw
	Caicai Zhang, PhD Associate Professor, The Hong Kong Polytechnic University caicai.zhang@polyu.edu.hk
	Yi-Fang Hsu PhD, National Taiwan Normal University yifanghsu@ntnu.edu.tw
	Laura Gwilliams, PhD New York University aura.gwilliams@nyu.edu
	Patrick Wang, PhD Professor, The Chinese University of Hong Kong p.wong@cuhk.edu.hk
Opposed Reviewers:	

Cover Letter

Professor Arturo Hernandez,
Editor-in-Chief of Journal of Neurolinguistics

April 14, 2023

Dear Professor Hernandez,

Accompanying this letter is the manuscript entitled "*Investigating MMN Responses to Lexical Tone Contrasts in Monolingual and Bilingual Speakers of Tonal Languages*", which I would like to submit for publication as a research paper in *Journal of Neurolinguistics*.

The present study aims to measure mismatch negativity (MMN) activity to tonal contrasts in monolingual and bilingual speakers of tonal languages to evaluate whether the language backgrounds of tonal language speakers influence their perceptual processes. The results showed that Hakka-Mandarin bilingual speakers showed MMN activity to tonal contrasts regardless of language type. This suggests that bilinguals have a more generalized representation of pitch heights that allows them to process the tonal information of different tonal languages. On the other hand, Mandarin monolingual speakers showed stronger MMN activity to Mandarin syllables than to Hakka syllables. This suggests that monolingual speakers may store a bank of phonological exemplars for perceiving the tonal information of native syllables. The results of this study can contribute to our understanding of how the brain processes and represents tonal information in different languages, which has implications for language learning and cross-linguistic communication. As one of the most influential and relied upon journals in the field of neuroscience, Journal of Neurolinguistics represents the perfect platform for us to share these results with the international research community.

The manuscript has been carefully proofread by a native English speaker and many grammatical errors have been corrected prior to submission. We confirm that this manuscript has not been published elsewhere and is not under consideration by another journal. All authors have approved the manuscript and agree with submission to Journal of Neurolinguistics. The authors have no conflicts of interest to declare.

Please address all correspondence to:
Dr. Chun-Hsien Hsu
Institute of Cognitive Neuroscience, National Central University, No. 300, Zhongda Rd., Zhongli District, Taoyuan City 320317, Taiwan
kevinhsu@ncu.edu.tw
Tel: 886-03-4227151 # 65207
Fax: 886-03-4263502

Yours sincerely,

Chun-Hsien Hsu, PhD
Assistant Professor
Institute of Cognitive Neuroscience
National Central University
https://deltaphase.github.io/Brain-and-Language-Lab-Eng/

Highlights

1. Hakka-Mandarin bilinguals showed similar MMN activity regardless of language type.
2. Monolinguals showed stronger MMN activity to native syllables
3. Bilingual speakers may have a generic tonal representation.

Abstract

The present study aimed to evaluate whether tonal language speakers' language backgrounds influence perceptual processes using an auditory oddball paradigm. Measurements of mismatch negativity (MMN) activity to tonal contrasts found that Hakka-Mandarin bilinguals showed similar MMN activity to tonal contrasts regardless of language type. Conversely, Mandarin monolingual speakers' data showed stronger MMN activity to Mandarin syllables than to Hakka syllables. The present study demonstrates that monolingual speakers might store a bank of phonological exemplars for perceiving the tonal information of native syllables, whereas bilinguals have a generic representation of pitch heights that allows them to process the tonal information of different tonal languages.

Keywords: MMN, lexical tones, tonal language, bilinguals

Title: Investigating MMN Responses to Lexical Tone Contrasts in Monolingual and Bilingual Speakers of Tonal Languages

Chun-Hsien Hsu[a], Wen-Chun Huang[b] and Tong-Hou Cheong[a]

[a] Institute Of Cognitive Neuroscience, National Central University, No. 300, Zhongda Rd., Zhongli District, Taoyuan City 320317, Taiwan

[b] Department of Hakka Language and Social Sciences, National Central University, No. 300, Zhongda Rd., Zhongli District, Taoyuan City 320317, Taiwan

Corresponding author: Chun-Hsien Hsu

Institute Of Cognitive Neuroscience, National Central University, No. 300, Zhongda Rd., Zhongli District, Taoyuan City 320317, Taiwan

kevinhsu@ncu.edu.tw

Tel: 886-03-4227151 # 65207

Fax: 886-03-4263502

Highlights

1. Hakka-Mandarin bilinguals showed similar MMN activity regardless of language type.
2. Monolinguals showed stronger MMN activity to native syllables
3. Bilingual speakers may have a generic tonal representation.

Abstract

The present study aimed to evaluate whether tonal language speakers' language backgrounds influence perceptual processes using an auditory oddball paradigm. Measurements of mismatch negativity (MMN) activity to tonal contrasts found that Hakka-Mandarin bilinguals showed similar MMN activity to tonal contrasts regardless of language type. Conversely, Mandarin monolingual speakers' data showed stronger MMN activity to Mandarin syllables than to Hakka syllables. The present study demonstrates that monolingual speakers might store a bank of phonological exemplars for perceiving the tonal information of native syllables, whereas bilinguals have a generic representation of pitch heights that allows them to process the tonal information of different tonal languages.

Keywords: MMN, lexical tones, tonal language, bilinguals

1. Introduction

The accurate perception of suprasegmental information is a crucial aspect of language processing, as it can provide listeners with important linguistic cues, such as affective-prosodic cues and prosodic phrasing. This is particularly relevant in tonal languages, such as Chinese and Thai, where pitch differences are used to differentiate words. While there have been studies of the role of pitch perception in both tonal and non-tonal languages (Chandrasekaran, Krishnan, & Gandour, 2007; T. Huang & Johnson, 2010), most research on language processing has taken a monolingual approach. However, given the increasing number of individuals living in multilingual environments, as by engaging with multilingual individuals and the media throughout their lives, it is essential to improve our understanding of the language processes of multilingual children and adolescents (Ayala, 2022; Bailey, Blackstock-Bernstein, & Heritage, 2015). In this context, researchers have emphasized the need for a greater understanding via work with different languages and dialects, as this may require specific needs (Garcia-Sierra, Ramirez-Esparza, Silva-Pereyra, Siard, & Champlin, 2012; Kuo, Uchikoshi, Kim, & Yang, 2016). However, this is not trivial. Some studies suggest that bilingual speakers use a single phonetic rule that works for both languages (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Williams, 1977), while others suggest that bilinguals may perceive identical acoustic-phonetic information differently depending on the language context (Elman, Diehl, & Buchwald, 1977; Garcia-Sierra et al., 2012). Nevertheless, it is essential to consider multilingualism in language processing research. Investigating the role of speech perception in a multilingual context can strengthen the theory of language acquisition and hold implications for spoken word recognition.

Kuo, Uchikosh, Kim, and Yang (2016) provided theoretical explanations for the impact of bilingual experiences on speech perception, such as the cross-language transfer theory and the structural sensitivity theory. According to the cross-language transfer theory, decoding and manipulating phonemic units in one language can be transferred to another language's phonetic competencies, but only if both languages share a linguistic feature and it is more prominent or complicated in one language than the other. On the other hand, the structural sensitivity theory suggests that speakers exposed to more than one language develop heightened sensitivity to the structural aspects of language and a greater readiness to recognize linguistic input to overcome interlingual interference (Kuo & Anderson, 2012).

This study does not aim to examine or provide definitive answers about the influence of bilingual experiences or to test these theories. Instead, it aims to explore whether the lexical tone contrasts of syllables are processed in the same way, regardless of the size of the tonal inventory of participants. This prediction was based on the notion that the size of the phonological inventory correlates with the ability to produce and perceive spoken sounds in monolinguals. For example, Teles and Huey (2020) demonstrated that speakers of languages with larger vowel inventories (e.g., English or French) would expand the space of vowel dispersion relative to speakers with smaller vowel inventories (e.g., Spanish), such that they would produce the target phonemes acoustically far away from one another.

Numerous studies of speech perception have focused on mismatch negativity (MMN), which is a unique event-related potential (ERP) used to detect distinguishable changes in acoustic features within a stream of sound and is not influenced by attention (Naatanen et al., 1997; Naatanen, Paavilainen, Tiitinen, Jiang, & Alho, 1993). The MMN paradigm typically involves a rapidly presented stream of repeated standard sounds occasionally interrupted by rare deviant sounds. MMN activity can be measured by comparing ERP responses to the deviant sound with those to the standard sound or by comparing ERP responses to the deviant sound in an MMN experiment with those to the same sound in an equal-probability control block (Jacobsen & Schroger, 2001). For example, native Japanese speakers often struggle to differentiate between the English phonemes /r/ and /l/, both of which are mapped to Japanese /l/. Zhang, Kuhl, Imada, Kotani, and Tohkura (2005) used magnetoencephalography (MEG) to record MMN in response to /r/ and /l/ sounds in native Japanese and native American English listeners. The study found that native Japanese listeners were less sensitive to the phonemic /r-l/ difference than native American English listeners, and their MMN amplitudes were significantly smaller. In the case of pitch perception, Mandarin Chinese syllables have four lexical tones: the high level tone (T1), high rising contour tone (T2), low falling-rising contour tone (T3), and high falling contour tone (T4). Chandrasekaran et al. (2007) used two experimental blocks with different tonal contrasts. In one block, the participants frequently heard the syllable /yi/ with T1 and occasionally heard /yi/ with T3. In the other block, the standard stimulus was /yi/ with T2. The results showed that native Mandarin speakers' MMN responses to the T1/T3 contrast were larger than their MMN responses to the T2/T3 contrast, indicating that MMN amplitudes are correlated with the acoustic similarity between pairs of standard and deviant sounds. Native English speakers' MMN did not demonstrate the effect of tonal contrast on MMN. Furthermore, while native Mandarin speakers' MMN to the T1/T3 contrast was larger than native English speakers' MMN to the T1/T3 contrast, there was no significant group difference in the T2/T3 contrast. These findings suggest that native tonal speakers are more sensitive to the height dimension than to the contour dimension, and that nontonal speakers' pitch perception does not appear to be significantly dependent on the height versus contour distinction.

As previously mentioned, little research has been conducted to evaluate the MMN responses of multilingual adolescents to tonal contrasts. To address this gap, the current study aimed to compare MMN responses to tonal contrasts in two groups of native speakers: Mandarin Chinese and Hakka-Mandarin Chinese bilingual speakers. Both languages used in this study are Sinitic languages. Hailu Hakka has a larger phoneme inventory than Mandarin Chinese, containing three tones in addition to the four of Mandarin Chinese. During the study, participants were presented with spoken syllables of one language in each experimental block. In the Mandarin block, the participants heard the Mandarin Chinese syllable /zu/ with T1 and T3, which are not real morphemes or words in Hailu Hakka. In the Hakka block, the participants heard the syllable /so/ with T1 and T3, which were not real morphemes or words in Mandarin either. It was anticipated that Hakka-Mandarin bilingual speakers would exhibit MMN responses to deviant stimuli in both the Mandarin and Hakka sessions. For Mandarin speakers, it was expected that their ERP activity would display MMN responses to tonal contrasts in the Mandarin session. In the Hakka blocks, if Mandarin speakers possess the generic ability to support their perception of pitch height, they may display MMN responses to Hakka syllables with T1/T3 contrast. However, owing to the simpler phonological structure of Mandarin Chinese than Hailu Hakka, cross-language transfer theory and structural sensitivity theory suggest that Mandarin speakers may be less sensitive to the phonemic features of Hakka syllables. Therefore, Mandarin speakers may exhibit reduced or insignificant MMN response to the T1/T3 contrast in Hakka syllables.

2. Methods

2.1 Participants

A total of 17 native Mandarin Chinese speakers and 16 Hakka-Mandarin bilinguals (aged 8–30 years) were recruited to participate in the MMN experiment. All the participants were college students with normal hearing and normal or corrected-to-normal vision. The current study was approved by the Human Subject Research Ethics Committee of National Taiwan University, Taiwan.

2.2 Stimuli

The experiment was conducted using two distinct sets of speech stimuli. The first set consisted of two Mandarin syllables, specifically /zu/ with T1 and T3, recorded by a male native Mandarin speaker. The second set of stimuli consisted of the Hakka syllable /so/ with T1 and T3, recorded by a male native Hailu Hakka speaker. Both T1 and T3 are lexical tones of Mandarin and Hailu Hakka. When recording these syllables, the native speakers were instructed to read a carrier sentence in their respective native languages with the target syllable placed at the end.

In addition, the selected Mandarin syllables are not words or morphemes in Hailu Hakka, and the selected Hakka syllables are not real words or morphemes in Mandarin. The syllables were then normalized to a duration of 350 ms and intensity of 70 dB using Praat. Speech waveforms and acoustic parameters of stimuli are listed in Table 1 and Figure 1.

2.3. Procedure

Jacobsen and Schröger (2001) suggested that a control procedure would allow control of the state of refractoriness during the oddball paradigm. Accordingly, each participant was required to undergo four experimental blocks in the present study, including the Mandarin-control block, Mandarin-MMN block, Hakka-control block, and Hakka-MMN block. Each experimental block comprised 500 trials. Each trial began with the presentation of a syllable lasting 350 ms (70 dB), followed by a 400-ms inter-trial interval. The syllables were presented using two loudspeakers. During the Mandarin sessions, participants heard Mandarin syllables /zu/, whereas during the Hakka sessions, they heard Hakka syllables /so/. In the control sessions, the T1 and T3 syllables were randomly presented with equal probabilities (p = 0.5), whereas in the MMN sessions, the T1 and T3 syllables were randomly presented in 100 (p = 0.2) and 400 (p = 0.8) trials, respectively. The order of the experimental blocks was counterbalanced across participants. While participating in the experiment, the participants watched a movie without sound or subtitles.

2.4. Data recording, preprocessing, and analysis

EEG data were recorded using 32 Ag/AgCl electrodes (QuickCap, Neuromedical Supplies, Sterling, USA). The electrodes were online-referenced to the average of the left and right mastoids for offline analysis. The EEG was continuously recorded and digitized at a rate of 1024 Hz, and the signal was amplified using a Grael 4 K EEG amplifier with a band-pass at DC–409 Hz. Electrode impedances were kept below 5 kΩ. Eye movements and blinks were monitored using supraorbital and infraorbital electrodes, and electrodes in the external canthi.

Instead of using conventional ERP analyses, we utilized the Hilbert-Huang transformation (HHT; Huang et al., 1998) for the offline analysis, because HHT can provide better resolution. In brief, HHT is a two-step method for analysis of nonlinear and nonstationary signals. The first step is empirical mode decomposition (EMD). EMD is a data-driven, adaptive method that decomposes a signal into a finite number of intrinsic mode functions (IMFs) determined using an iterative sifting process that separates the signal into high and low frequency components. This decomposition method is adaptive and can automatically adjust to the signal's frequency content. Once the signal is decomposed into IMFs, the instantaneous phase of each IMF can be calculated using the direct quadrature transform. The instantaneous frequency can then be obtained by calculating the time derivative of the instantaneous phase. This approach allows for better frequency resolution as it estimates the frequency content of the signal at each time point, rather than using methods such as the convolution integral method used in Fourier transform and wavelet transform. Therefore, HHT is useful in the analysis of non-linear and non-stationary signals. In the present study, EEG data were analyzed in the following manner based on the procedure described by Hsu, Lee, & Liang, (2016). Continuous EEG data were epoched with 100-ms pre-stimulus intervals and 600-ms post-stimulus intervals. The pre-stimulus interval (−100 to 0 ms) was used for baseline correction. Trials were rejected if they were contaminated by voltage variations larger than ±100 μV. We then decomposed each EEG segment into seven IMFs using the masked empirical mode decomposition (Quinn, Lopes-Dos-Santos, Dupret, Nobre, & Woolrich, 2021) and obtained event-related modes (ERMs) by averaging IMFs across trials (Al-Subari, Al-Baddai, Tome, Goldhacker, et al., 2015; Al-Subari, Al-Baddai, Tome, Volberg, et al., 2015; Chen, Chao, Chang, Hsu, & Lee, 2016). Similar to noise removal using filters, ERMs were obtained by summing the IMFs based on their instantaneous frequencies and then averaging across trials. The correlation between MMN components and frequency bands is well known, and Kalyakin et al. (2007) has suggested that EEG activity between 2 and 8 Hz reflects MMN. Previous studies of MMN responses using the

HHT method confirmed that EMD can be used to analyze EEG signals and employ IMFs with frequencies ranging between 2 and 8 Hz to estimate MMN activity (Cong et al., 2009; Hsu et al., 2016). Therefore, this study focused on IMFs with frequencies between 2 and 8 Hz to extract MMN-related activities.

2.4.1. Statistical Analysis

In subsequent statistical analyses, the MMN effect was measured by comparing the ERPs to the deviant stimuli of the MMN clock and the same stimuli of the equal probability control block. The grand-averaged waveforms for T1 syllables in the MMN and control blocks are presented in Figure 2. Through visual inspection of the data, a negative-going component peaking at around 300 ms was observed, and this component was more prominent in the waveforms of the MMN blocks than in those of the control blocks. We used mass univariate cluster-based permutation tests as recommended by Maris and Oostenveld (2007) to evaluate the significance of MMN activity of each participant group and each syllable. This cluster-based nonparametric approach is recommended to control Type I error rates in electrophysiology experiments where precise latencies and scalp distribution are unknown a priori. The mass univariate cluster-based permutation tests was run using the Eelbrain package (http://doi.org/10.5281/zenodo.3923991). The general procedure for the cluster-based test was as follows: A paired t-statistic (deviant minus control, one tail) was calculated at each time point and channels. Spatiotemporal clusters were then formed from test statistics that were contiguously significant (uncorrected p = 0.01) through time and channels. For each cluster, the cluster mass statistic was computed, which was the sum of all t values in the cluster. To determine the reliability of these clusters, the actually observed cluster-level test statistics were compared against the null distribution based on 10,000 random permutations of the condition labels.

In addition, we also analyze the MMN waveforms obtained by subtracting ERPs to a T1 stimuli of the control blocks from those of the MMN blocks. Then, the mass univariate cluster-based permutation tests were applied to evaluate the effect of language types (MMN of Mandarin syllables minus MMN of Hakka syllables) and the effect of participant groups (Hakka-Mandarin bilinguals minus Mandarin speakers). In the analysis of MMN waveforms, spatiotemporal clusters were formed from test statistics that were contiguously significant (uncorrected p = 0.1) through time and channels. This criterion could increase the sensitivity to weak and longer-lasting effects and did not affect the false alarm rate of the non-parametric statistical test.

3. Results

3.1. Analysis of MMN effect (deviant minus control) within each group

The finding among Mandarin speakers demonstrated significant MMN effects at two spatiotemporal clusters (Figure 3A). This suggests that the deviant stimuli in the MMN blocks produced more negative activity than those in the control blocks. For Mandarin speakers' ERPs to Mandarin syllables /zu1/, the significant spatiotemporal cluster ($p = .001$) ranged from 260 to 366 ms and was located in 23 electrodes, including frontal, central, and posterior electrodes. On the other hand, Mandarin speakers' ERPs to Hakka syllables showed a significant cluster ($p = .011$) ranged from 292 to 354 ms, and this cluster was located in 21 electrodes, including frontal, central, and left-posterior electrodes.

For Hakka-Mandarin bilinguals (Figure 3B), the statistical analysis for ERPs to Mandarin syllables yielded a significant spatiotemporal cluster ($p = 0.0005$), which ranged from 224 to 375 ms located in 23 electrodes, including frontal, central, and posterior electrodes. Lastly, Hakka-Mandarin bilinguals' ERP to Hakka syllables also showed a significant cluster ($p = .035$) ranged from 278 to 331 ms, and this cluster was located in 14 electrodes, including frontal-anterior, right-anterior and right-central electrodes.

3.2. Analysis of MMN waveforms

For the effect of language types (Figure 3C), Mandarin speakers showed a significant cluster ($p = 0.021$) which ranged from 237 to 365 ms and was located in almost all electrodes except FT7, T3 and TP7. Hakka-Mandarin bilinguals' difference waveforms did not show the effect of language types. Finally, the analyses of the participant groups failed to show any significant cluster.

4. Conclusion

In summary, the present study found that each participant group showed MMN activity to tonal contrast regardless of language type. The analysis of the MMN waveforms demonstrated that Mandarin speakers showed stronger MMN activity to Mandarin syllables than to Hakka syllables. The theoretical implications of these results are as follows. The ERP measures of both participant groups showed MMN activity to both languages. This finding implies that establishing perception skills for pitch height in syllable perception may be essential for speakers of tonal language. This speculation was supported by the MEG study of Chinese as a second language (CSL) learners reported by Lee et al. (2017), which demonstrated that CSL learners' MMN activity to the T1/T3 contrast generated in the left hemisphere was enhanced after three days of language training, but their MMN to the T2/T3 contrast did not change after training. Nevertheless, future studies of language acquisition are needed to explore whether the sensitivity to pitch heights and contours superimposed on spoken syllables can predict language acquisition ability in multilingual environments.

Note that the phoneme and tonal inventories of Hailu Hakka are greater than those of Mandarin Chinese. One might therefore expect that the MMN of Hakka-Mandarin bilinguals would be larger than that of Mandarin monolinguals, because previous studies of speech production and ERPs have demonstrated the effect of phonological inventory size on vowel dispersion and MMN activity. Specifically, Hacquard et al.'s (2007) ERP study demonstrated that language users with a large phoneme inventory (French) showed facilitation in perceiving the contrast of speech sounds compared to those with a small inventory (Spanish). The correlation between the size of the vowel inventory and the magnitude of MMN activity is associated with the findings of Teles and Huey (2020), who noted that speakers of languages with larger vowel inventories (e.g., English or French) produce target phonemes acoustically dissimilar to one another and show an expansion of the space of vowel dispersion relative to speakers with smaller vowel inventories (e.g., Spanish). Although our results did not show a significant difference between Hakka-Mandarin bilinguals and Mandarin monolinguals in amplitudes, the present result showed a differences in the onset latency of spatiotemporal clusters. That is, Hakka-Mandarin bilinguals showed spatiotemporal clusters of MMN activity (onsets of spatiotemporal clusters of MMN effect to Mandarin syllables: 224 ms; to Hakka syllables: 278 ms) slightly earlier than that of Mandarin speakers (onsets of spatiotemporal clusters of

MMN effect to Mandarin syllables: 260 ms; onsets of Hakka syllable: 292 ms). It seems to imply that the automatic detection mechanism of tonal changes is more sufficient in Hakka-Mandarin bilinguals than in Mandarin speakers. Nevertheless, although the size of the tonal inventory might different across language, the size of the tonal space seems not to be different across tonal languages. A quantitative investigation of production data recorded from nine tonal languages, which showed a fixed pitch range across tonal languages, regardless of the size of the tone inventory (Kuang, 2013). A further implication of the present finding is that although the phonological inventory account can predict interlanguage differences in speech perception, this account seems likely to reflect the effect of formant frequency dispersion on speech perception. More future studies of tonal perception are need to evaluate whether participants' F0 ranges and tonal inventory would influence pitch perception performance.

In addition, bilingual studies have emphasized that bilinguals tends to generate fewer words than monolinguals in vocabulary fluency tasks (Bialystok, Craik, & Luk, 2008; Gollan, Montoya, & Werner, 2002; Portocarrero, Burright, & Donovick, 2007; Rosselli et al., 2000) in which participants were asked to report as many as possible of a group of words of the same semantic category or the same initial letters. This is because bilinguals may have a smaller overall vocabulary than monolinguals in each language and in each must deal with competition from the other language. In this regard, one might ask whether the perceptual processes in bilinguals are less efficient than those in monolinguals. Interestingly, the present study showed that the MMN activities of Hakka-Mandarin bilinguals in both languages were similar to those of Mandarin monolinguals to Mandarin syllables. Our findings imply that tonal language speakers' perception of changes in pitch height superimposed on familiar syllables is not sensitive to participants' lexical fluency.

Finally, the present study found that Mandarin speakers exhibited MMN activity to Hakka syllable /so/ which is not a real word or real morpheme in Mandarin Chinese. This effect seems to be consistent with the notion mentioned earlier that tonal speakers are sensitive to the height of lexical tones. One might wonder that would it raised a question to the exemplar theory of phonology (Goldinger, 1996, 1998). According to exemplar theory, the mental representation of a linguistic category includes episodic traces of spoken words. Note that the syllable /zu/ is a real morpheme in Mandarin but an accidental gap in Hailu Hakka, and the syllable /so/ is a real morpheme in Hailu Hakka but an accidental gap in Mandarin Chinese. Therefore, the exemplar theory would predict the absence of MMN activity to Hakka syllables among Mandarin speakers, because Mandarin monolinguals would have an inability to perceive the pitch height of non-native syllables. Interestingly, the present data demonstrated that Mandarin speakers showed stronger MMN in the hearing of Mandarin syllables than in the hearing of Hakka syllable. On the other hand, Hakka-Mandarin bilinguals showed similar MMN responses to both languages. Therefore, to some extent, the present data can still be explained by the exemplar theory. That is, although Mandarin speakers may have a more generalized representation of pitch heights for lexical tones, their MMN to non-native syllables was reduced comparing to that to native syllables. In future studies, it would be interesting to see how nonnative speakers acquire syllables and lexical tones and manifest the phonological representations of native speakers in tonal languages.

In conclusion, the present study demonstrates that both monolinguals and bilinguals to tonal languages might store a generic representation to perceive the tonal information of syllables, whereas bilinguals have a more efficient mechanism of pitch height processing that allows them to process tonal information from different tonal languages. This study highlights the importance of considering the linguistic background of the participants when studying speech perception in tonal languages, and underscores the need for more research to better understand how linguistic experience can be used to improve language learning.

5. Acknowledgements

This work was supported by the National Science and Technology Council of

Taiwan (grant number: 112-2628-H-008 -002), and the Ministry of Education of

Taiwan.

References

Al-Subari, K., Al-Baddai, S., Tome, A. M., Goldhacker, M., Faltermeier, R., & Lang, E. W. (2015). EMDLAB: A toolbox for analysis of single-trial EEG dynamics using empirical mode decomposition. *J Neurosci Methods, 253*, 193-205. doi:10.1016/j.jneumeth.2015.06.020

Al-Subari, K., Al-Baddai, S., Tome, A. M., Volberg, G., Hammwohner, R., & Lang, E. W. (2015). Ensemble Empirical Mode Decomposition Analysis of EEG Data Collected during a Contour Integration Task. *PLoS One, 10*(4), e0119489. doi:10.1371/journal.pone.0119489

Ayala, A. G. (2022). At the intersection of culture, bilingualism, and mathematics: breaking language norms in a seventh-grade dual-language classroom for biliteracy development. *International Journal of Bilingual Education and Bilingualism, 25*(6), 2109-2123. doi:10.1080/13670050.2020.1859984

Bailey, A. L., Blackstock-Bernstein, A., & Heritage, M. (2015). At the intersection of mathematics and language: Examining mathematical strategies and explanations by grade and English learner status. *The Journal of Mathematical Behavior, 40*, 6-28.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J Exp Psychol Learn Mem Cogn, 34*(4), 859-873. doi:10.1037/0278-7393.34.4.859

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *J Acoust Soc Am, 54*(2), 421-428. doi:10.1121/1.1913594

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

Chen, W. F., Chao, P. C., Chang, Y. N., Hsu, C. H., & Lee, C. Y. (2016). Effects of orthographic consistency and homophone density on Chinese spoken word recognition. *Brain and Language, 157-158*, 51-62. doi:10.1016/j.bandl.2016.04.005

Cong, F., Sipola, T., Huttunen-Scott, T., Xu, X., Ristaniemi, T., & Lyytinen, H. (2009). Hilbert-Huang versus Morlet wavelet transformation on mismatch negativity of children in uninterrupted sound paradigm. *Nonlinear Biomed Phys, 3*(1), 1. doi:10.1186/1753-4631-3-1

Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual Switching in Bilinguals. *Journal of the Acoustical Society of America, 62*(4), 971-974. doi:Doi 10.1121/1.381591

Garcia-Sierra, A., Ramirez-Esparza, N., Silva-Pereyra, J., Siard, J., & Champlin, C. A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain and Language, 121*(3), 194-205. doi:10.1016/j.bandl.2012.03.008

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *J Exp Psychol Learn Mem Cogn, 22*(5), 1166-1183. doi:10.1037//0278-7393.22.5.1166

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychol Rev, 105*(2), 251-279. doi:10.1037/0033-295x.105.2.251

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16*(4), 562-576. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12382994

Hacquard, V., Walter, M. A., & Marantz, A. (2007). The effects of inventory on vowel perception in French and Spanish: an MEG study. *Brain and Language, 100*(3), 295-300. doi:10.1016/j.bandl.2006.04.009

Hsu, C. H., Lee, C. Y., & Liang, W. K. (2016). An improved method for measuring mismatch negativity using ensemble empirical mode decomposition. *Journal of Neuroscience Methods, 264*, 78-85. doi:10.1016/j.jneumeth.2016.02.015

Huang, N. E., Shen, Z., Long, S. R., Wu, M. L. C., Shih, H. H., Zheng, Q. N., . . . Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences, 454*(1971), 903-995. doi:DOI 10.1098/rspa.1998.0193

Huang, T., & Johnson, K. (2010). Language Specificity in Speech Perception: Perception of Mandarin Tones by Native and Nonnative Listeners. *Phonetica, 67*(4), 243-267. doi:10.1159/000327392

Jacobsen, T., & Schroger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology, 38*(4), 723-727. doi:Doi 10.1111/1469-8986.3840723

Kalyakin, I., Gonzalez, N., Joutsensalo, J., Huttunen, T., Kaartinen, J., & Lyytinen, H. (2007). Optimal digital filtering versus difference waves on the mismatch negativity in an uninterrupted sound paradigm. *Dev Neuropsychol, 31*(3), 429-452. doi:10.1080/87565640701229607

Kuang, J. (2013). The tonal space of contrastive five level tones. *Phonetica, 70*(1-2), 1-23. doi:10.1159/000353853

Kuo, L. J., & Anderson, R. C. (2012). Effects of early bilingualism on learning phonological regularities in a new language. *Journal of Experimental Child Psychology, 111*(3), 455-467. doi:10.1016/j.jecp.2011.08.013

Kuo, L. J., Uchikoshi, Y., Kim, T. J., & Yang, X. Y. (2016). Bilingualism and phonological awareness: Re-examining theories of cross-language transfer and structural sensitivity. *Contemporary Educational Psychology, 46*, 1-9. doi:10.1016/j.cedpsych.2016.03.002

Lee, R. R. W., Hsu, C. H., Lin, S. K., Wu, D. H., & Tzeng, O. J. L. (2017). Learning transforms functional organization for Mandarin lexical tone discrimination in the brain: Evidence from a MEG experiment on second language learning. *Journal of Neurolinguistics, 42*, 124-139.

Naatanen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., . . . Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*(6615), 432-434. doi:DOI 10.1038/385432a0

Naatanen, R., Paavilainen, P., Tiitinen, H., Jiang, D., & Alho, K. (1993). Attention and Mismatch Negativity. *Psychophysiology, 30*(5), 436-450. doi:DOI 10.1111/j.1469-8986.1993.tb02067.x

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Arch Clin Neuropsychol, 22*(3), 415-422. doi:10.1016/j.acn.2007.01.015

Quinn, A. J., Lopes-Dos-Santos, V., Dupret, D., Nobre, A. C., & Woolrich, M. W. (2021). EMD: Empirical Mode Decomposition and Hilbert-Huang Spectral Analyses in Python. *J Open Source Softw, 6*(59). doi:10.21105/joss.02977

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solis, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Appl Neuropsychol, 7*(1), 17-24. doi:10.1207/S15324826AN0701_3

Teles, L., & Huey, O. (2020). The Effects of the French Vowel Inventory on Vowel Production in Spanish Speakers. *University of Rochester working papers in the language sciences, 8*(1), 15-31.

Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics, 21*(4), 289-297.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *Neuroimage, 26*(3), 703-720. doi:10.1016/j.neuroimage.2005.02.040

Legend

Table 1. Voice onset time (VOT) and the first three formant frequencies for each stimuli.

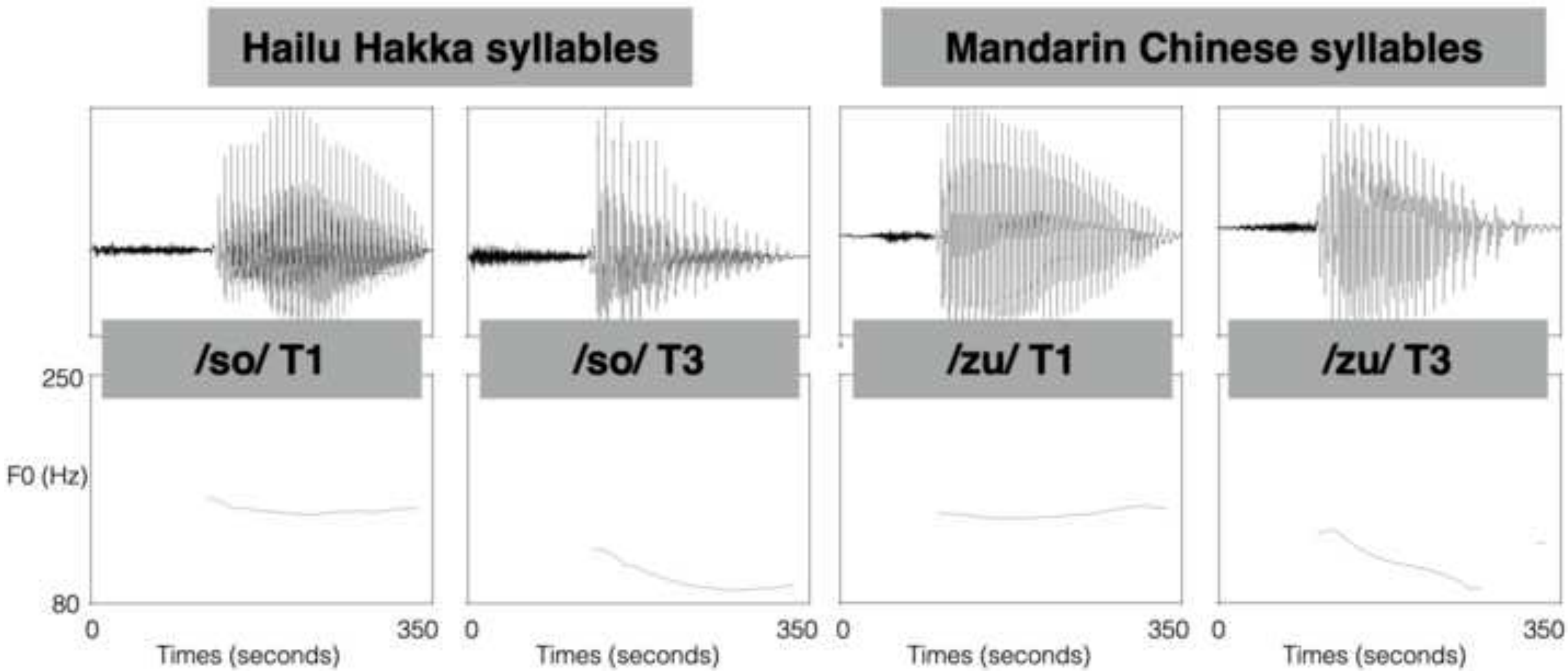
Figure 1. Speech waveforms and F0 contours of stimuli.

Figure 2. Grand average ERPs elicited by syllables in MMN and control blocks within each participant group at electrodes of interest.

Figure 3. (A) MMN effect (deviant minus control) of Mandarin participants in the hearing of Mandarin syllables and Hakka syllables. The outlined region is significantly different between deviant and control stimuli based on a mass-univariate t-test. (B) MMN effect (deviant minus control) of Hakka-Mnadarin participants. (C) Language effect of MMN difference waveforms (Mandarin syllables minus Hakka Syllables) of Mandarin participants.

Figure1

Click here to access/download;Figure;Figure1.png



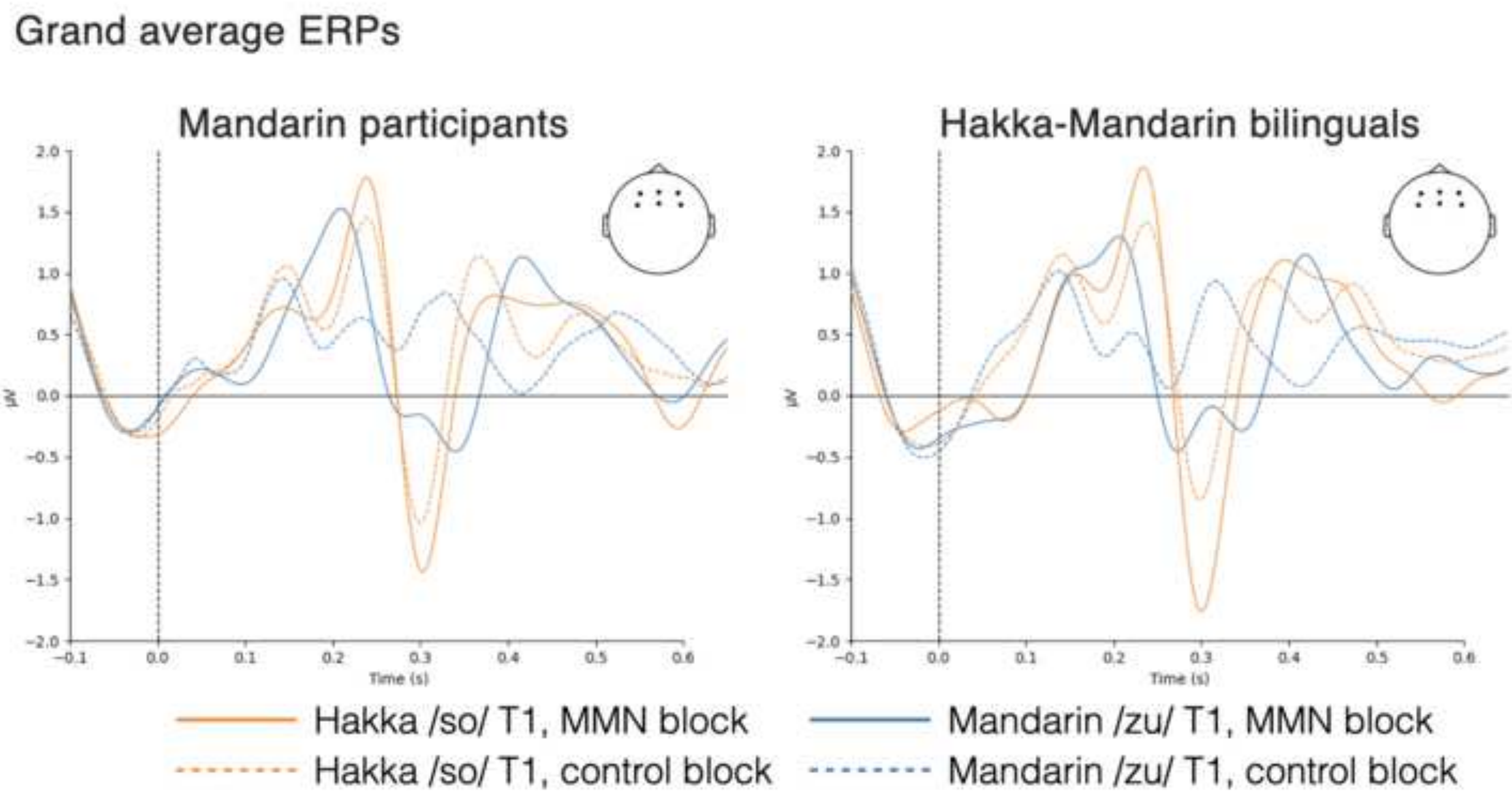

Grand average ERPs
Mandarin participants
Hakka-Mandarin bilinguals
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
μV
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Time (s)
Hakka /so/ T1, MMN block
Hakka /so/ T1, control block
Mandarin /zu/ T1, MMN block
Mandarin /zu/ T1, control block

Figure3

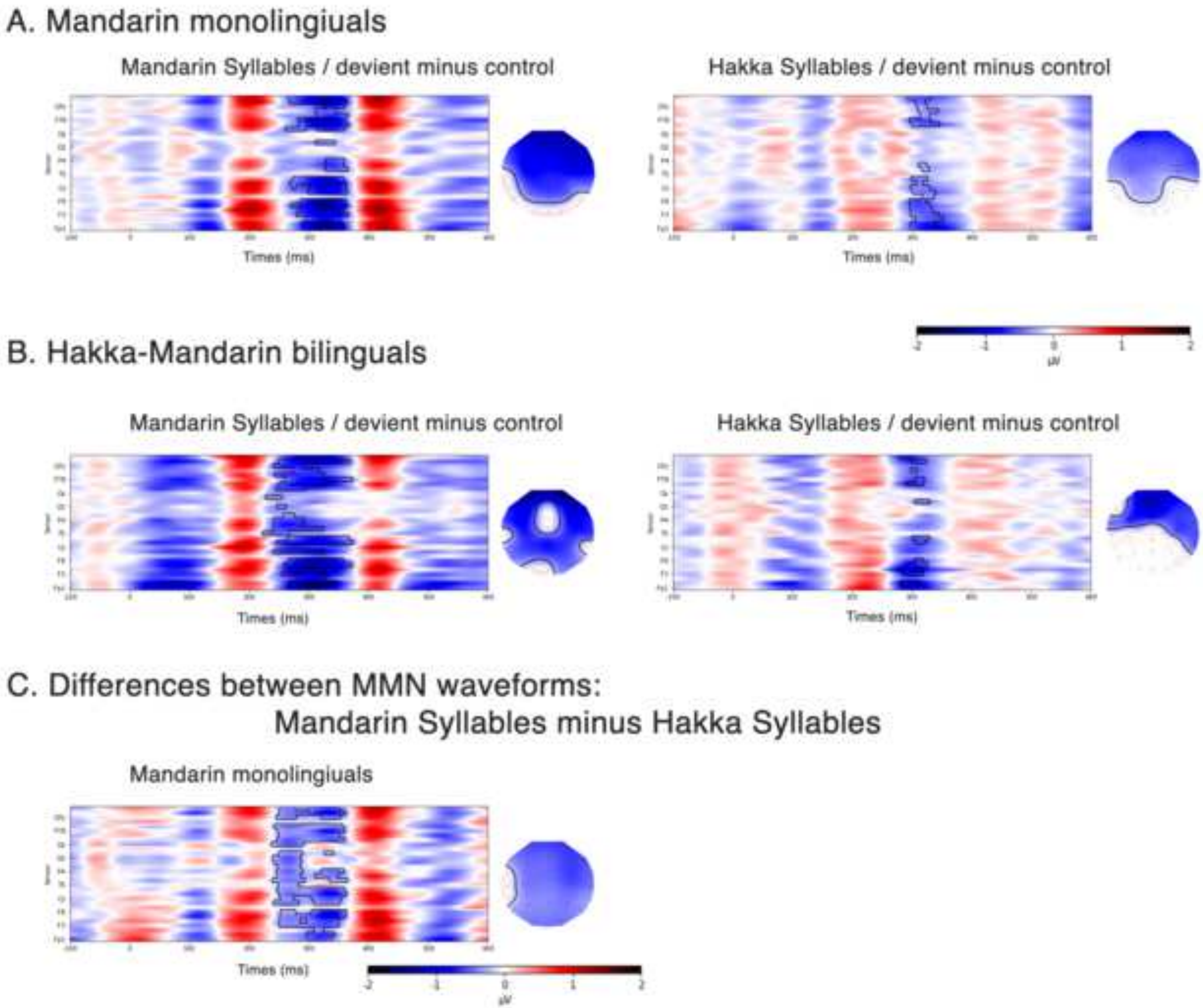

A. Mandarin monolingiuals
Mandarin Syllables / devient minus control
Times (ms)
Hakka Syllables / devient minus control
Times (ms)
B. Hakka-Mandarin bilinguals
μV
Mandarin Syllables / devient minus control
Times (ms)
Hakka Syllables / devient minus control
Times (ms)
C. Differences between MMN waveforms:
Mandarin Syllables minus Hakka Syllables
Mandarin monolingiuals
Times (ms)
μV

Table 1. Voice onset time (VOT) and the first three formant frequencies for each stimuli.

	Hakka /so/ T1	Hakka /so/ T3	Mandarin /zu/ T1	Mandarin /zu/ T3
VOT (ms)	122	122	94	99
F1 (Hz)	624	645	402	393
F2 (Hz)	1030	1059	851	764
F3 (Hz)	2815	2813	2962	2691

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: