Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Pre-attentive Perceptual Integration of Tones and Vowels

William Choi [1], *Xiuli Tong*[2]

[1, 2] Division of Speech and Hearing Sciences, The University of Hong Kong, HKSAR
choitm@hku.hk, xltong@hku.hk

Abstract

The feature-integration theory of attention posits that auditory features are first processed independently at early stages. We tested the theory by examining the neural processing integrality of vowels and tones at the early pre-attentive auditory level. Twenty native Cantonese listeners participated in the event-related potential (ERP) experiment. We adopted the mismatch negativity (MMN) additivity approach, and elicited three types of mismatch negativities (tone deviant, vowel deviant and double-tone-vowel deviant) using a passive oddball paradigm. We found that the mismatch negativity of the double deviant was not additive by the mismatch negativities of the single deviants. The results indicate that tones and vowels are processed integrally by the same neuronal population at the early pre-attentive auditory level. Potential implications to theories and models of speech perception will be discussed.

Index Terms: vowel, tone, perception, integral processing, mismatch negativity, mismatch negativity additivity

1. Introduction

To date, how auditory information is processed in human brains remains unclear. According to the feature-integration theory of attention (Treisman & Gelade, 1980), auditory information were first differentiated and processed independently at the early stage, then processed in an integral manner at later stages under attentional control. Such a theory has received support from neurophysiological studies which suggested the independent processing of various acoustic features, e.g., intensity-fundamental frequency (Paavilainen, Valppu, & Näätänen , 2001; Takegata, Paavilainen, Näätänen , & Winkler, 1999) and duration-fundamental frequency (Levanen, Hari, McEvoy, & Sams, 1993). Adopting the same approach, a recent study suggested the integral processing of consonants and tones among Chinese listeners (Gao, Hu, Gong, Chen, Kendrick, & Yao, 2012). It remains unclear whether tones are integrally processed with another important segmental information, i.e., vowels. In order to provide a full picture for speech perception among tonal listeners, the current study set out to examine the perceptual integrality of vowels and tones at the early pre-attentive auditory level.

The mismatch negativity additivity approach has been proven to be a valid tool for testing the feature-integration theory of attention (Treisman & Gelade, 1980). The approach is based on the additivity of mismatch negativity, an event-related potential component which reflects pre-attentive neuronal detection of sound change. In this approach, two single deviant and one double deviant mismatch negativities are required. Using Paavilainen et al. (2001) to illustrate, two single deviant Mismatch negativities, i.e., fundamental frequency-mismatch negativity and intensity-mismatch negativity were elicited using a passive oddball paradigm. On top of the two single deviant mismatch negativities, a double deviant mismatch negativity was elicited with a deviant stimulus which differed from the standard stimulus in both fundamental frequency and intensity. If the two acoustic features are independently processed, in this case fundamental frequency and intensity, the double deviant mismatch negativity should be equal to the sum of the two single deviant mismatch negativities. This notion has been supported by localization studies (Caclin et al., 2006; Takegata et al., 2001) confirming that mismatch negativity additivity reflects separate mismatch negativity generators and distinct auditory memory traces for each acoustic feature. On the other hand, if the two acoustic features are integrally processed by the same neuronal population, the double deviant mismatch negativity would not be equal to the sum of the single deviant mismatch negativities. According to the neuronal under firing account, when a neuron reacts to two acoustic features, the intensity of neuronal firing is either equal to or slightly higher than when it reacts to single deviant changes (Lidji, Jolicoeur, Kolinsky, Moreau, Connolly, & Peretz, 2010).

Early neurophysiological studies adopting this mismatch negativity additivity approach offered support to the feature integration theory of attention (Paavilainen et al., 2001; Levanen et al., 1993; Takegata et al., 1999). Paavilainen et al. elicited three types of mismatch negativities among adult listeners – fundamental frequency mismatch negativity, intensity mismatch negativity, and double mismatch negativity. The amplitudes of the single deviant mismatch negativities were summed, and no significant difference was found between the summed amplitude and the double mismatch negativity amplitude. A similar result was reported in Levanen et al.'s study which investigated the perceptual integrality of

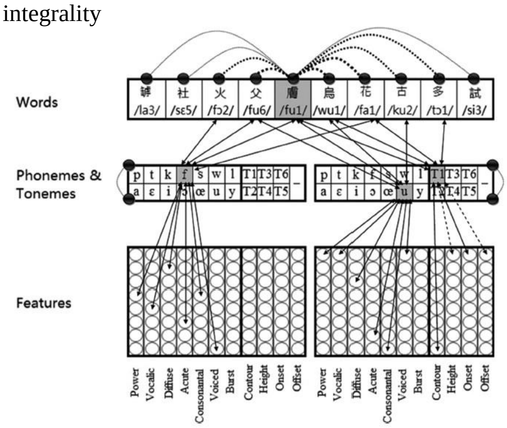

Figure 1. *Pictorial representation of the TTRACE Model. The tonal and segmental information are integrally processed by same processing units at auditory (feature) and phonological (phonemic and tonemic) levels. Adopted from Tong et al. (2014).*

 doi: 10.21437/SpeechProsody.2016-65

fundamental frequency and duration. Together, the results suggested that fundamental frequency, intensity and duration were independently processed by separate neuronal populations at the pre-attentive auditory level. This was consistent with the feature integration theory
of attention which posited that acoustic features were processed independently at low level. Worthy to note, the studies above only tested acoustic features, i.e., fundamental frequency, duration and intensity, which did not have a lexical role in the non-tonal population they tested. Of interest in the current study is the acoustic features which have prominent statuses in making lexical distinctions, such as tones and vowels among tonal populations.

A recent study has offered preliminary support to pre-attentive tone-vowel perceptual integrality. In that study, the single mismatch negativities of vowel and fundamental frequency deviants were non-additive, suggesting that vowels and pitch were integrally processed by the same neuronal population (Lidji et al., 2010). However, the results must remain tentative as the acoustic parameters of the stimuli were not strictly controlled. For instance, the pitch deviant stimulus had marked differences from the standard stimulus not only in fundamental, but also formant frequencies. Stimuli aside, the study examined French listeners whose language did not make use of pitch for lexical distinctions. Worthy to note, 70% of the world's languages entail the lexical use of pitch (Yip, 2002). However, little is known about the perceptual integrality of tone and segmental information. This question is of high theoretical significance as segmental-tone integrality (or non-integrality) is a fundamental claim in the speech perception models which incorporate both segmental and tonal processing among tonal listeners (Tong, McBride-Chang, & Burnham, 2014; Ye & Connine, 2002).

Specifically, the TTRACE model adopted the integral view of segmental and tonal processing (Tong, McBride, & Burnham, 2014). Different from the original TRACE model (McClelland & Elman, 1986), the TTRACE model has a new tonal layer, and incorporates the integral processing of consonants, vowels and tones at both phonological and auditory levels (see Figure 1). However, differentiating the level at which processing integrality occurs, namely phonological and auditory, is beyond the scope of behavioral data (Holder, 1992; Law, Fung, & Kung, 2013). In view of this, the current study set out to examine the neuronal processing integrality of tones and vowels. Additionally, we aim to unfold the temporal processes and examine whether the processing integrality occurs as early as the pre-attentive auditory level.

In short, the current study set out to (a) examine the neural processing integrality of tone and vowel and (b) evaluate whether the integrality occurs at the early pre-attentive auditory level among tonal listeners. To serve this purpose, the current study has adopted the mismatch negativity additivity approach and evaluated the neural perception of vowel, tone, and vowel-tone double deviants among native Cantonese listeners. We hypothesize that tone and vowels are integrally processed at the early pre-attentive auditory level, to be indicated by the non-additivity of tone and vowel mismatch negativities.

2. Method

2.1. Participants

A total of 20 native Cantonese right handers (9 males and 11 females) were recruited from Hong Kong. Their mean age was 22 years 4 months (SD = 3 years 11 months). All participants passed a pure tone hearing screening.

2.2. Stimuli and procedures

Four phonologically legal pseudowords /u1/ /i1/ /u3/ /i3/ were artificially synthesized using the source-filter synthesis technique. Each stimulus was used reversibly at different blocks as standard and deviant stimuli. We strictly controlled the acoustic parameters of the stimuli, such that the vowel contrasts, i.e., /u1/-/i1/ and /u3/-/i3/ differed within themselves only in formant frequency. The tonal contrasts, i.e., /u1/-/u3/ and /i1/-/i3/ differed only in fundamental frequency. The double contrasts, i.e., /u1/-/i3/ and /i1/-/u3/, differed in both fundamental frequency and formant frequency (see Table 1). We chose tone 1 and tone 3 since they were both level tones, allowing strict fundamental frequency height manipulations. This set of tone contrasts has been used in previous behavioral and neurophysiological studies (Francis et al., 2003; Gu et al., 2013).

During online electroencephalographic (EEG) recording, participants were instructed to sit comfortably and watch a silent movie, while being presented with the stimuli audibly by E-Prime 2.2 via inserted earphones. The stimulus-onset asynchrony was set to 900ms with reference to a previous adult mismatch negativity study on Cantonese tone perception (Gu et al., 2013).

The whole experiment contained three oddball and three reverse oddball blocks. The order of the blocks was pseudo-randomized. In the three oddball blocks, /u1/ was used throughout as the standard stimulus, while /i1/, /u3/ and /i3/ served as deviants. In the reverse oddball blocks, /u1/ was the deviant throughout, while /i1/, /u3/ and /i3/ served as standard stimuli. Each block consisted of 60 deviant and 540 standard stimuli. The order of the stimuli was pseudo-randomized in a way that adjacent deviant stimuli were separated by at least two standard stimuli. Altogether, there were 360 deviant and 3240 standard stimuli.

2.3. EEG recording and offline processing

The electroencephalogram was recorded with a 64-channel silver chloride (Ag-AgCl) Quik-Cap with electrodes placed on 64 scalp sites following the International 10-20 system. The Quik-Cap was then connected to a SynAmps 2 amplifier
designed to detect very small traces of electrical signals under the scalp. The left and right mastoids served as the reference electrodes. We collected the electrical signals at a sampling rate of 500 Hz with the bandpass of 0.05-70 Hz. The signal were then amplified by the amplifier. The impedances of all electrodes were kept below 5kΩ throughout.
The electroencephalographic data were processed and analyzed offline with NeuroScan 4.5 system. The electroencephalographic signals were first offline filtered with a bandpass of 1 to 30 Hz. Subsequently, the resulting data were segmented into epochs from 100 ms before onset to 600 ms post-onset of the stimuli (Gu et al., 2013). The two standard stimuli following each deviant stimulus were not analyzed due to the possible mismatch negativity effect of

Table 1. *Acoustic parameters of Cantonese tone stimuli*

	F0 (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	Dur. (ms)	Int. (dB)
u1	140	312	827	3k	4k	300	70
u3	113	312	827	3k	4k	300	70
i1	140	312	2223	3k	4k	300	70
i3	113	312	2223	3k	4k	300	70

deviant stimulus on the following standard stimuli. After baseline correction, we removed artifacts beyond -80μV or +80μV. After that, we obtained the average event-related potential of each stimulus type by averaging the grand event-related potential of the respective stimulus type.

3. Results

We set the mismatch negativity time window between 150 to 250 ms after the onset of deviant stimulus (Chandrasekaran et al., 2007; Tong et al., 2014; Tsang et al., 2010). Grand-averaged waves inspection showed the largest mismatch negativity amplitudes at the frontal-central (Fz) electrode relative to other frontal electrodes, consistent with previous mismatch negativity studies (Gu et al., 2013; Näätänen et al., 2007; Tong et al., 2014). Thus, the Mismatch negativities we report below are based on the subtraction of standard stimuli's event-related potential signal from that of the deviant stimuli at the Fz electrode. Using the same criteria in Gu et al.'s study, we obtained the average amplitudes of the Mismatch negativities at a time frame of 40ms, 20ms before and after the highest point of the grand averaged mismatch negativity peak. The ERP waveforms are shown in Figure 2.

We computed a new variable "summed mismatch negativity" by summing up the mean mismatch negativity amplitudes of the single deviants, i.e., tone deviant and vowel deviant. Next, we conducted paired-sample t-tests to evaluate whether the amplitude of the "summed mismatch negativity" was significantly different from the mismatch negativity amplitude of the double deviant. Paired sample t-tests revealed a significant difference between the "summed mismatch negativity" and the mismatch negativity of double deviant, $t(19) = 3.05$, $p < .01$. This reflects that the mismatch negativity of the double deviant, i.e., vowel-tone deviant, is non-additive by the Mismatch negativities of single deviants.

4. Discussion

In this study, we (a) investigated whether tones and vowels were integrally or independently processed and (b) examined whether the processing integrality occurred at the early pre-attentive auditory level among tonal listeners. Statistical analysis revealed significant difference between the mismatch negativity of the double deviant and the summed mismatch negativity of the single deviants. The mismatch negativity non-additivity suggests the integral processing of tones and vowels by the same neuronal population. Additionally, the processing integrality occurs at the early pre-attentive auditory level.

The processing integrality of tones and vowels found in the current study is consistent with previous neural (Lidji et al., 2010) and psycholinguistic studies (Lee & Nusbaum, 1993; Repp & Lin, 1990; Tong et al., 2008; Tong et al., 2014). On top of this, we went beyond behavioral findings in the temporal dimension, by probing into an early stage of perceptual processing, i.e., pre-attentive auditory processing, and found the recruitment of the same neuronal population for tone and vowel processing. Worthy to note, this finding is in contrary to the feature-integration theory of attention (Treisman & Gelade, 1980) which postulated that auditory information were first processed independently at the early stage, and integrally processed at later attentional stages.

Although the theory was reinforced by similar mismatch negativity additivity studies which demonstrated independent processing of acoustic features such as intensity-fundamental frequency (Paavilainen, Valppu, & Näätänen , 2001; Takegata, Paavilainen, Näätänen , & Winkler, 1999;) and duration-fundamental frequency (Levanen, Hari, McEvoy, & Sams, 1993) at pre-attentive auditory stage, the early integral vowel-tone processing found herein indicates that early independent processing of auditory features is not necessarily the case.

Tonal language experience may offer an account for the early integrality of tone-vowel processing. In terms of lexical function, vowels and tones are important dimensions which tonal language listeners attend to for lexical identification. On the other hand, intensity and duration had very little role, if any, in lexical identification. Thus, it could be the case that tonal listeners' perceptual systems were adapted at stages early as the pre-attentive auditory stage where the same neuronal population was recruited for the formant frequency (vowels) and fundamental frequency (tones) processing. This notion is supported by previous neurophysiological evidences demonstrating that pre-attentive auditory processing can be shaped by tonal language experience. For instance, while the processing of pitch is lateralized at the right hemisphere among non-tonal listeners (Jamison, Watkins, Bishop, & Matthews, 2006; Okamoto, Stracke, Draganova, & Pantev, 2009), tonal language listeners demonstrated left hemispheric processing of pitch information (Gu, Zhang, Hu, & Zhao, 2013). Additionally, Krishnan, Gandour, and Bidelman (2010) revealed that pitch tracking sensitivity at the peripheral cochlear level could be enhanced by long-term tonal language experience. Taken together, the above studies point to the same notion that low levels of auditory processing are susceptible to shaping by language experience. Admittedly, the role of language experience in shaping the early processing integrality of formant and fundamental frequencies remains tentative as the current study lacks a control group, i.e., non-tonal listeners. A new study with strict acoustic manipulations showing independent processing of formant frequency and fundamental frequency among non-tonal listeners would give further support to this plausible account. Nevertheless, the current finding provides a new insight to traditional speech processing theories – not all acoustic dimensions are

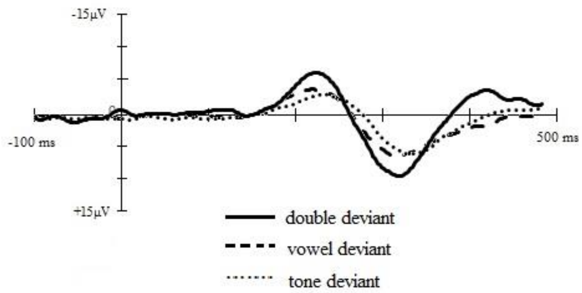

Figure 2. *Deviant minus standard ERP waveforms under double deviant, vowel deviant and tone deviant conditions at Fz electrode*

independently processed at the auditory level; and that early pre-attentive auditory processing might be attuned by language experience.

The early integral processing of tone and vowel found herein is in line with the prediction of the TTRACE model (see Figure 2). Adopting the same paradigm, a recent study found the non-additivity of consonant and tone Mismatch negativities, indicating the integral processing of consonants and tones among Chinese listeners (Gao et al., 2012). Notably, an early tone perception model posited the tonal layer as a separate representation overlaid upon segmental layer (Ye & Connine, 1999). However, the current findings and those of Gao et al. suggest that segmental information (consonants and vowels) and tones are not independently but integrally processed by the same neuronal population. These are inconsistent with the separate representation view of Ye and Connine, but in line with the fundamental claim of the TTRACE - integral processing of segmental and suprasegmental information. In addition, we have extended this claim in the temporal domain, by providing neural evidence that the processing integrality occurs early at the pre-attentive auditory level.

The findings might also offer potential directions in revising speech perception models other than the TTRACE. In the C-Cure Model (McMurray & Jongman, 2011), tonal listeners might form hypothesis of the average fundamental frequency, then compare the acoustic input with the expected average fundamental frequency. The expectation is then refined for the identification of the next tone and revision of the previous tone activated. Applying the current findings to this model, the processing of suprasegmental cues may interact with the processing of acoustic cues related to segmental information at the early auditory level. Similarly, a new suprasegmental layer could be added to the Distributed Cohort Model (Gaskell & Marslen-Wilson, 1997). During online speech perception and lexical processing, the conditional probability of the lexical candidate may be determined by a holistic rather than separate evaluation of segmental and suprasegmental information. Future studies may consider testing the feasibility of adding a suprasegmental layer in the existing models by means of computational modeling.

Apart from its theoretical and methodological significances, the current study offers potential directions for future studies to examine segmental and suprasegmental processing integrality. First, we evaluated only level tone contrasts but not contour tones. Interestingly, behavioral evidence did not seem to support processing integrality of contour tones and segmental information (Tong et al., 2014). Future studies may examine a whole set of tone contrasts in examination of this issue. Additionally, only a limited set of vowels have been tested in the current study. Apart from that, a large number of consonantal properties such as frication and voice onset time have not been explored. We encourage future studies to explore different acoustic variations to further enhance the external validity in the segmental and suprasegmental interaction so found in the current study.

Admittedly, the current findings do not solve some inherent problems of the TTRACE model. In the TRACE (and TTRACE) model, all mental lexicons are actively evaluated during hypothesis testing (McClelland & Elman, 1986; Tong et al., 2014). In this framework, Norris (1994) estimated that the processing of a six-phoneme word in a language with 50000 lexicons would require 550000 lexical processing units; and 113 billion connections resulting from bidirectional activation and inhibition within and across layers. With such an enormous and unrealistic number of connections required for the processing of only one word, the TRACE and TTRACE models might need further refinement.

5. Conclusions

In conclusion, we provide neurophysiological evidence on the perceptual integrality of tone and vowels. On top of this, we extended previous psycholinguistic studies in the temporal domain, by suggesting that the perceptual integrality occurred at the early pre-attentive auditory level. Apart from its theoretical contributions to models and theories of speech perception, the current study has a methodological contribution, by using a new neurophysiological paradigm to examine the perceptual integrality of segmental and suprasegmental information among tonal populations.

6. References

[1] Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Experience-dependent neural plasticity is sensitive to shape of pitch contours. *Cognitive Neuroscience and Neuropsychology*, *18*(18), 1963-1967.

[2] Caclin, A., Brattico, E., Tervaniemi, M., Näätänen , R., Morlet, D., Giard, M. H., & McAdams, S. (2006). Separate neural processing of timbre dimensions in auditory sensory memory. *Journal of Cognitive Neuroscience*, *18*(12), 1959-1972.

[3] Francis, A., Ciocca, V., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics*, *65*(7), 1029-1044.

[4] 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. doi:10.1371/journal.pone.0038289

[5] Garner, W. R. (1974). *The processing of information and structure*. New York: Wiley.

[6] Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*, 613-656.

[7] Gu, F., Zhang, C., Hu, A., & Zhao, G. (2013). Left hemisphere lateralization for lexical and acoustic pitch processing in Cantonese speakers as revealed by mismatch negativity. *Neuroimage* (*83*), 637-645.

[8] Holder, D. S. (1992). Electrical Impedance Tomography of brain function. *Brain topography*, *5*(2), 87-93

[9] Jamison, H. L., Watkins, K. E., Bishop, D. V., & Matthews, P. M. (2006). Hemispheric specialization for processing auditory nonspeech stimuli. *Cerebral Cortex*, *16*(9), 1266-1275.

[10] Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, *23*, 81-95.

[11] Law, S., Fung, R., & Kung, C. (2013). An ERP study of good production vis-a-vis poor perception of tones in Cantonese: Implications for top-down speech processing. *PLoS ONE*, *8*(1): e54396. doi:10.1371/journal.pone.0054396

[12] Lee, L., & Nusbaum, H. C. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception & Psychophysics*, *53*, 157–165.

[13] Levänen, S., Hari, R., McEvoy, L., Sams, M. (1993) Responses of the human auditory cortex to changes in one versus two stimulus features. *Experimental Brain Research*, *97*, 177-183.

[14] Lidji, P., Jolicœur, P., Kolinsky, R., Moreau, P., Connolly, J. F., Peretz, I. (2010). Early integration of vowel and pitch processing: A mismatch negativity study. *Clinical Neurophysiology*, *121*, 533-541.

[15] McClelland, J. L., & Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology*, *18*, 1-86.

[16] McMurray, B., & Jongman, A. (2011). What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review*. Advance online publication. doi:10.1037/a0022325

[17] Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., & Foxe. J.J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, *14*, 115-129.

[18] Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (mismatch negativity) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590.

[19] Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189-234.

[20] Okamoto, H., Stracke, H., Draganova, R., & Pantev, C. (2009). Hemispheric asymmetry of auditory evoked fields elicited by spectral versus temporal stimulus change. *Cerebral Cortex*, *19*(10), 2290-2297.

[21] Paavilainen, P., Mikkonen, M., Kilpelainen, M., Lehtinen, R., Saarela, M., & Tapola, L. (2003). Evidence for the different additivity of the temporal and frontal generators of mismatch negativity: A human auditory event-related potential study. *Neuroscience Letters*, *349*(2), 79-82.

[22] 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 approximately 2 but 1+1+1<3. *Neuroscience Letters*, *301*(3), 179-182.

[23] Repp, B. H., & Lin, H.-B. (1990). Integration of segmental and tonal information in speech perception: A cross-linguistic study. *Journal of Phonetics*, *18*, 481-495.

[24] Takegata, R., Huotilainen, M., Rinne, T., Näätänen , R., & Winkler, I. (2001). Changes in acoustic features and their conjunctions are processed by separate neuronal populations. *NeuroReport*, *12*, 525-529.

[25] Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neuroscience Letters*, *266*, 109-112.

[26] 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.

[27] Tong, X., McBride, C., Lee, C. Y., Zhang, J., Shuai, L., Maurer, U., & Chung, K. K. H. (2014). Segmental and suprasegmental features in speech perception in Cantonese-speaking second graders: An ERP study. *Psychophysiology*, *51*(11), 1158-1168.

[28] Tong, X., McBride-Chang, 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*, 1589-1605.

[29] Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97-136.

[30] Tsang, Y., Jia, S., Huang, J., & Chen, H. (2010). ERP correlates of pre-attentive processing of Cantonese lexical tones: The effects of pitch contour and pitch height. *Neuroscience Letters*, *487*, 268-272.

[31] Watkins,S., Dalton,P., Lavie,N., & Rees,G. (2007). Brain mechanisms mediating auditory attentional capture in humans. *Cerebral Cortex*, *17*(7), 1694-1700.

[32] Wolff C, & Schröger E (2001) Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. *Neuroscience Letters*, *311*, 37-40.

[33] Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes*, *14*, 906-630.

[34] Yip, M. (2002). *Tone*. Cambridge textbooks in linguistics. Cambridge: Cambridge University Press.

[35] Ylinen S., Huotilainen M., & Näätänen, R. (2005). Phoneme quality and quantity are processed independently in the human brain. *NeuroReport*, *16*, 1857-1860.