Journal of Neurolinguistics 41 (2017) 11–23

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Research paper

On the early neural perceptual integrality of tones and vowels

William Choi [a], Xiuli Tong [a,*], Feng Gu [a], Xiuhong Tong [b], Lena Wong [a]

[a] Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong
[b] Institute of Psychological Sciences, Zhejiang Key Laboratory for Research in Assessment of Cognitive Impairments, and Center for Cognition and Brain Disorders, Hangzhou Normal University, China

ARTICLE INFO

Article history:
Received 1 January 2016
Received in revised form 30 August 2016
Accepted 6 September 2016
Available online 14 September 2016

Keywords:
Vowel
Tone
Perception
Integral processing
MMN
MMN additivity
TTRACE model

ABSTRACT

The current study adopted the MMN additivity approach to examine the pre-attentive perceptual integration of vowels and tones. Twenty Cantonese listeners participated in the ERP experiment. Using the passive oddball paradigm, we elicited tone-MMN, vowel-MMN and double-MMN in the speech condition; and fundamental frequency-MMN, formant frequency-MMN and double-MMN in the non-speech condition. In both conditions, the double-MMNs were significantly smaller in amplitude than the sum of single feature MMNs. Morphological comparisons showed no significant difference in the latency and topographic patterns between vowel-MMN and tone-MMN, and marginal significant differences between formant frequency-MMN and fundamental frequency-MMN. Collectively, results reflect the perceptual integration of tones and vowels at the phonological level, and partial integration of fundamental frequency and formant frequency at the auditory level.

Integrality, by definition, is the conjunctive processing of features (Kolinsky, Lidji, Peretz, Besson, & Morais, 2009; Sussman, Gomes, Nousak, Ritter, & Vaughan, 1998). To date, there is an ongoing debate on whether segmental and tonal information are integrally or independently processed. This question is theoretically important as it is a fundamental assumption among the two models for tone perception, i.e., the modified TRACE model (Ye & Connine, 1999) and the TTRACE model (Tong, McBride-Chang, & Burnham, 2014). Taking the independent view, the modified TRACE assumed that the tonal layer was a separate representation overlaid upon the segmental layer. They posited that segmental and tonal information were processed independently by separate processing units, without any perceptual interaction. On the other hand, the TTRACE model took the opposite view, and incorporated a tonal layer integrated with that of segmental information. In the TTRACE model, a tonal listener first conjunctively analyzes features of tones (e.g., fundamental frequency contour and height) and segmental information (e.g., burst, vocalic and frication). The multiple activation of features results in the mutual activation of phonemes (e.g., /f/ /u/) and tones (e.g., high level tone/T1, mid level tone/T3) at the phonological level. The mutual activation of phonemes and tones would then activate syllables (which are also words in Chinese) (e.g., /fu1/ /wu1/ /si3/) at the lexical level, in which the target lexicon receives the highest level of activation (see Fig. 1). According to the TTRACE model, at the auditory and phonological levels of speech perception, consonants, vowels and tones are integrally processed by the same processing units.

The integral processing of segmental and tonal information received support from psycholinguistic studies which looked into the Garner effect or Garner interference. The Garner interference is an increase in reaction time caused by the inclusion of

* Corresponding author.
E-mail address: xltong@hku.hk (X. Tong).

http://dx.doi.org/10.1016/j.jneuroling.2016.09.003
0911-6044/

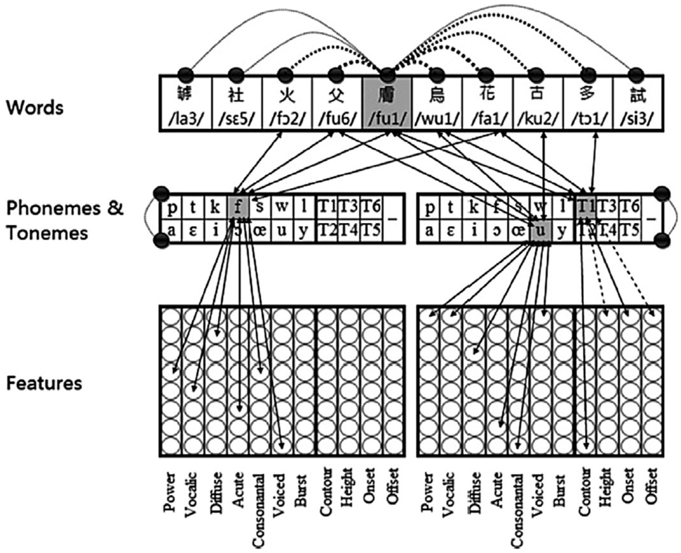

Fig. 1. Pictorial representation of the TTRACE Model. The tonal and segmental information are integrally processed by same processing units at auditory (feature) and phonological (phonemic and tonemic) levels. Adopted from Tong et al. (2014).

irrelevant information during perceptual processing (Garner, 1974). For example, Repp and Lin (1990) adopted a speeded classification paradigm in which the listeners were asked to classify consonants (/b/ or /d/), vowels (/a/ or /u/) and tones (high or falling). Taking the tone classification task as an example, the two tones were carried by the same syllable (e.g., /ba/) in the control condition; whilst the two tones were carried by different syllables containing varying vowel information (e.g., /ba/ and /bu/) in the orthogonal condition. In both conditions, participants were asked to selectively attend to the tones, while ignoring the irrelevant information, namely consonants and vowels. It was assumed that if the irrelevant information was processed integrally with the relevant information, the inclusion of the variation of irrelevant information in the orthogonal condition would increase the reaction time. Indeed, the results indicated that tonal listeners had longer reaction times under the orthogonal than control conditions. The observation of Garner interference between tones and segments suggests the integral processing of tonal and segmental information (Lee & Nusbaum, 1993; Repp & Lin, 1990; Tong, Francis, & Gandour, 2008).

However, although the Garner interference has frequently been claimed to reflect perceptual integrality, it may not be a strictly valid measure of such. In fact, increased memory workload may account for the Garner interference (Janczyk & Kunde, 2012). As illustrated in the previous paragraph, the syllables under the control condition only differed in one dimension (e.g., tone) while those under the orthogonal condition differed in two dimensions (e.g., vowel and tone). Being asked overtly to ignore the irrelevant dimension, the participants had to divert cognitive resources to inhibit or suppress the irrelevant dimension. Hence, fewer cognitive resources were left to process the relevant dimension, resulting in longer reaction time in the orthogonal than control conditions. This view is paralleled by the unexpected results observed in two studies using Garner paradigm (Lee & Nusbaum, 1993; Repp & Lin, 1990), in which the English listeners also showed Garner interferences in the segmental and tonal classification tasks. With the absence of lexical tones in English, it is unreasonable that non-tonal English listeners would process lexical tones with segmental information integrally. Thus, the universality of "perceptual integrality" so found might be suggestive of an alternate explanation of the Garner interferences - increased cognitive demands rather than language specific integral processing.

In a recent year, Tong et al. (2014)'s study offered new evidence to the perceptual integrality of segmental and tonal information. In an odd-one-out tone discrimination task, the participants were asked to point out the syllable that carrying a different tone among four syllables. The segmental information of all the four syllables varied systematically across trials in three conditions. In one condition, they were made up of the same consonant but different vowels (e.g. /ŋɐi5/ /ŋɔ4/ /ŋɐu4/ /ŋai4/). In the second condition, they were made up of different consonants but with the same vowels (e.g. /tsiu1/ /jiu1/ /siu3/ /biu1/). In the third condition they were made up of different consonants and different vowels (e.g. /hɐi5/ /ji5/ /niu5/ /tsœŋ6/). It was found that the perceptual accuracy of odd-one-out discrimination task varied across the three conditions, suggesting a perceptual interaction between segmental and tonal information. However, the authors also provided an alternative account for the observed segmental context effects on tone perception - differential masking effect of formant frequencies on fundamental frequency perception. Given the above validity concerns and those in the Garner paradigm, psycholinguistic evidences can only offer limited support to the perceptual integrality of segmental and tonal information.

Notably, neurophysiological technique offers a more direct way to investigate the mechanism of speech perception by probing into cortical activities (Näätänen et al., 1997). Of interest in the current study is an event related potential (ERP) component called mismatch negativity (MMN). MMN is generated by neuronal activities in response to sound changes, and