Perceptual attunement to coarticulation: hearing tone in vowel height

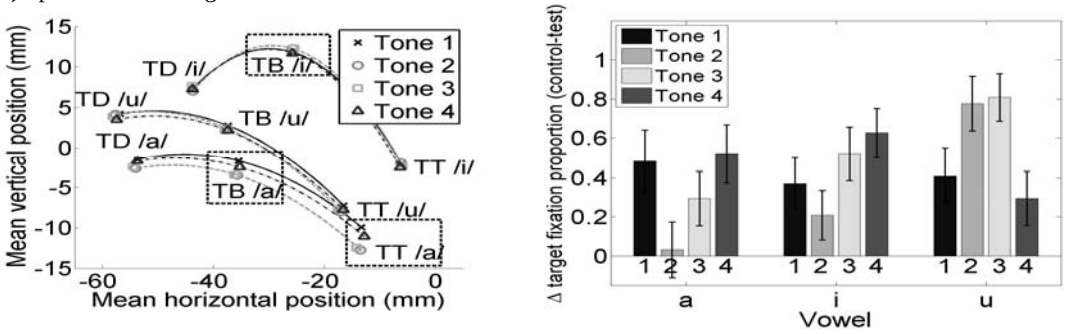
Jason A. Shaw[1], *Wei-rong Chen*[2], *Michael D. Tyler*[1], *Donald Derrick*[1], *Michael Proctor*[3]
[1]University of Western Sydney, [2]National Tsing Hua University, [3]Macquarie University
j.shaw@uws.edu.au, fjueng@yahoo.com.tw, M.Tyler@uws.edu.au, D.Derrick@uws.edu.au, michael.proctor@mq.edu.au

From the perspectives of phonological theory (e.g., Yip, 2002) and of the source-filter theory of speech production (e.g., Fant, 1960), tones and vowels are largely, if not entirely, independent. With rare (and controversial) exception, tones do not condition sound changes or synchronic alternations in vowels. However, a few articulatory studies have found that lexical tone production in Mandarin Chinese exerts an influence on tongue height in vowels, particularly on low vowels (Hoole and Fu, 2004; Erickson, *et al.*, 2004). In complementary production (Electromagnetic Articulography, EMA) and perception (eye-tracking in the visual world paradigm) experiments, we replicate and extend past results on tone-vowel coarticulation and then demonstrate the relevance of these results to speech perception. We find that *the influence of tone on vowel height in production facilitates early recognition of tone in perception.*

Method (production): We recorded six native speakers of Mandarin Chinese (3 male) using the NDI WAVE EMA system. Sensors were attached to the tongue tip (TT), blade (TB), dorsum (TD), lips, jaw, nasion and mastoids. Each speaker produced /pi/, /pa/ and /pu/ syllables with each of the four Mandarin tones: 1 *high*, 2 *low-high*, 3 *low*, and 4 *high-low*. Each syllable was produced 12 times by each speaker, generating a corpus of 864 tokens (12 reps x 3 vowels x 4 tones x 6 speakers). Syllables were presented in Pinyin and randomized with fillers. After computationally correcting for head movement, vowel targets were determined by a 20% threshold of peak velocity of the TD sensor in the opening movement of the vowel. Fig. 1a reports tongue position for each tone-vowel combination.

Method (perception): To investigate the time course of tone perception, we used eye-tracking in a printed-word version of the visual world paradigm. In the test condition, four monosyllabic words differing only in tone, e.g., 八 /pa1/, 拔 /pa2/, 把 /pa3/, 爸 /pa4/, were displayed in Chinese characters while one of the words, e.g., /pa2/, was played over a speaker. Target words combined the four Mandarin tones with three vowels: /i/, /a/, /u/. Sixteen native speakers of Mandarin Chinese were asked to click on the word that they heard, while a Tobii x120 eye-tracker recorded fixations to words displayed in the four corners of the screen. Fixations to the target word in the test condition were compared to fixations in a control condition where competitor words were phonologically dissimilar to the target, e.g., target /pa2/ amongst competitors /ti1/, /t^{h}a3/, /tu4/. Fig. 1b reports difference in fixation proportion to target words across control and test conditions. Smaller differences across conditions (smaller bars) indicate faster tone recognition.

Figure 1: (left panel) Tongue position for tone-vowel combinations. Dotted line rectangles indicate significant effects of tone **(right panel)** Speed of tone recognition across vowels. Smaller bars indicate faster tone identification. Error bars indicate 95% CI.



Results: As indicated by the dotted rectangles in Fig. 1a, significant effects of tone on vowel position were found at the TB sensor for both /i/ and /a/ (and at the TT sensor for /a/ only). Tone 2 lowered the tongue blade for /a/ and raised the tongue blade for /i/. The vowel /u/ was stable across tones. In perception, the speed of tone recognition varied across vowels. For /u/, tones 1 and 4 (tones that start high) were recognized fastest, potentially due to the higher intrinsic F0 of /u/ (Whalen and Levitt, 1995). For /a/ and /i/, tone 2, the tone that exerted a reliable coarticulatory influence on tongue height, was recognized faster than the other tones. Moreover, /a2/, which showed the largest effect in articulation was recognized faster than any other tone-vowel combination.

Discussion: Misattribution of the source of coarticulation is argued to be a major factor in sound change (Ohala, 1981). Although our EMA results show consistent coarticulatory influences of tone on vowel height (for certain vowels), these have not led historically to tone-conditioned vowel alternations. Our perception results indicate consistent attribution of tongue height variation to tone. This ability to perceive tone in tongue height may stem from the intrinsically coextensive nature of tones and vowels and contribute to the apparent resistance of tone-conditioned vowel change.