On-line processing of voice pitch feedback during production of Mandarin tones

Yi Xu, Charles R. Larson, and Jay J. Bauer
Department of Communication Sciences and Disorders
Northwestern University
Evanston, IL 60208

An important question in studying speech production is whether articulations are the result of pre-planned movement patterns or whether on-line sensory feedback is used to help guide the movements. Many previous studies addressing this issue have returned conflicting results. One extreme theoretical position is that every detail of an articulatory movement is planned out well before execution, so that little on-line adjustment is needed. At the other extreme one may argue that planning goes only as far as some canonical form of the underlying articulatory target, and the implementation of the underlying targets is done with continual guidance and correction directed by sensory feedback. The pitch-shift auditory feedback paradigm developed in recent years (Burnett et al., 1998; Hain et al., 2000; Larson et al., 1995) presents a convenient tool for providing answers to questions concerning articulatory planning and the role of sensory feedback in speech production. The present study investigates the effect of pitch-shifted auditory feedback on the production of lexical tones in Mandarin disyllabic words. Five native speakers of Mandarin Chinese were recorded saying four different bitonal patterns without any carrier frame in two experimental conditions each with 80 repetitions. In one condition, for approximately one half the trials, the voice pitch feedback (over headphones) was shifted up 200 cents (200 ms duration) beginning 170 ms after onset of vocalization. In the other condition the voice pitch was shifted down 200 ms. The timing of the pitch-shift stimulus was such that it started during the first Mandarin syllable and ended during the second syllable. Preliminary data analysis from two subjects indicated that, although the speakers produced similar bitonal F_0 patterns as previously reported (Xu, 1997), they compensated for the shift in pitch by either raising (for downward pitch-shifts) or lowering (for upward pitch-shifts) their voice fundamental frequency (F0) beginning about 130 ms after stimulus onset and lasting for about 200 ms. The response magnitude was about 50 cents. These results are similar to those seen during sustained vowel phonations in previous studies (Burnett et al., 1998; Hain et al., 2000), which were interpreted to be reflexive responses to perturbed auditory pitch feedback. The results from the present study indicate that on-line monitoring of auditory feedback and F0 adjustment are also essential components of syllable productions in Mandarin. Recognizing that auditory feedback is but one of at least three major feedback channels (two others being tactile and proprioceptive), these data may be considered as the speaker's reflexive adjustment in reaction to the discrepancy between auditory feedback and the combined tactile and proprioceptive feedback. These data may also indicate that when a speaker's F0 does not reach a desired level, task-dependent, compensatory, reflexive mechanisms may help to achieve and maintain the appropriate F0 level. Thus, auditory feedback is used by speech mechanisms to assist the speaker in producing articulatory goals necessary for a particular linguistic expression.

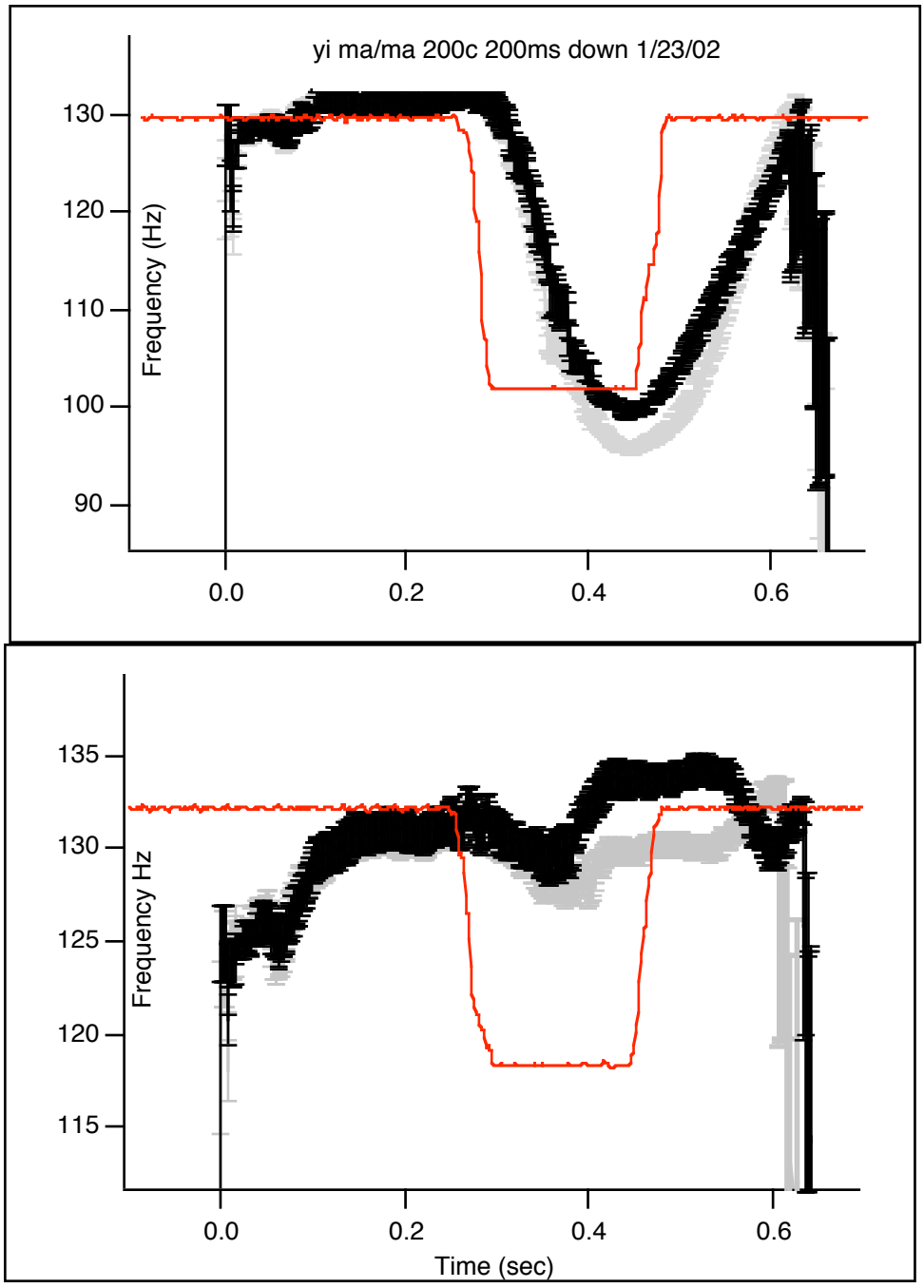
References:

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (**1998**). "Voice f0 responses to manipulations in pitch feedback," J. Acoust. Soc. Am. **103**, 3153-3161.

Hain, T. C., Larson, C. R., Burnett, T. A., Kiran, S., and Singh, S. (**2000**). "Instructing participants to make a voluntary response reveals the presence of two vocal responses to pitch-shift stimuli," Exp. Brain Res. **130**, 133-141.

Larson, C. R., Carrell, T. D., Senner, J. E., Burnett, T. A., and Nichols, L. L. (**1995**). "A proposal for the study of voice fo control using the pitch shifting technique," in *Vocal fold physiology : Voice quality control,* edited by O. Fujimura and M. Hirano (Singular, San Diego), 321-331.

Xu, Y. (**1997**). "Contextual tonal variations in Mandarin," *Journal of Phonetics* **25**, 61-83.



Examples of responses to shifted feedback (black) and control (grey). In each example, the subject produced the syllables "ma ma". The top set shows responses in which the second syllable has a falling then rising F0 contour. The bottom set shows responses in which the second syllable has a relatively steady F0 contour. Each set of traces is an average of 15 responses. Width of black and grey traces represents standard error of the mean Superimposed on response traces is a trace representing the timing of the pitch-shift stimulus, in this case a downward pitch-shift.