Prosodic encoding of information structure in Mandarin:
An experimental investigation of a tone language

Iris Chuoying Ouyang and Elsi Kaiser
University of Southern California

Prosody conveys discourse-level information [1][2], but the extent to which prosodic cues distinguish different kinds of information remains unclear ([2][3] on new-information/contrastive focus). The prosodic encoding of discourse-information is even more complicated in *tone languages*, where cues such as F0, duration, and intensity also distinguish lexical items (e.g. Mandarin Chinese [4][5]). Prior work in Mandarin led to divergent findings regarding (i) presence/absence of differences between focus-types (new vs. corrective/contrastive) and (ii) what cues—if any—mark differences between focus-types (lengthening, F0 range expansion, intensity), e.g.[6][7].

We conducted a **production study** on Mandarin to investigate whether **(i)** the presence/absence of correction and **(ii)** the new/given distinction are encoded prosodically (2x2, 36 targets, 36 fillers). Participants (4 women, 4 men) produced instructions to move an object to a location, indicated by pictures and arrows (Fig.1), e.g. '*Move the* $\underline{bamboo}_{TARGET}$ *next to the fridge*' (Ex.1). Target nouns were bisyllabic, with High-High, High-Low, Low-High tonal contours (HH/HL/LH).

In **Non-corrective** conditions, the correct object (e.g. bamboo) moved after the command. In **Corrective** conditions, an incorrect object moved after the participant first produced the command (e.g. the *sunglasses* moved next to the fridge)—so participants had to *repeat* the command to correct the incorrect movement. In **New** conditions, the target word had not been mentioned on that trial until the participant first used it. In **Given** conditions, the target word occurred in a correct movement earlier on the same trial (i.e., known to hearer). There were four conditions: Corr|New, Corr|Given, NoCorr|New, NoCorr|Given.

Results-The ***presence/absence of correction*** was reflected in all three acoustic parameters: Target words in Corrective conditions had *longer durations, and larger F0- and intensity-ranges* than Non-corrective words (p's<.03).

The ***new/given distinction*** was reflected in duration and F0, but only in Non-Corrective conditions (significant Correction-by-Givenness interaction). ***Corrective*** conditions had no differences between given/new. In ***Non-corrective*** conditions, New information had longer duration, larger F0-ranges than Given information (p's<.04); intensity-ranges did not differ. This suggests new-information focus is encoded differently from correction: Only correction is associated with expanded intensity-ranges.

The absence of given/new effects in Corrective conditions may be due to acoustic marking of givenness/newness being defined from the speaker's perspective: In all Corrective conditions, the target words had already been *uttered by the speaker,* although the hearer apparently had not heard them properly the first time.

Discussion-Even in a language with lexical tones, which differ in F0, duration, and intensity, all of these acoustic dimensions also encode information-structure. We found no evidence for a 'specialized-function' situation where some cues mark information-structure and others mark lexical distinctions. Instead, *all three dimensions* are multi-functional. However, further analyses showed that intensity-range expansion was due to *minimum-intensity lowering*, while F0-range expansion resulted from *maximum-F0 raising*, pointing to pressures/constraints on how information-structure is encoded. Combining our work with prior claims, it seems that variations in the *ranges* of *both F0 and intensity* can mark information-structure, whereas their *shapes/contours* carry information about lexical tones [4][5]. This highlights the fine-grained ability of the language production system to utilize different aspects of acoustic dimensions.

(1) ba OBJECT fangdao LOCATION pangbian '*Move the OBJECT next to the LOCATION*'
(particle) OBJECT put LOCATION side

Fig. 1: Sample display of stimuli

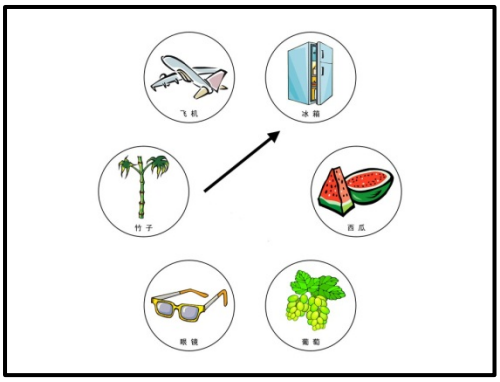
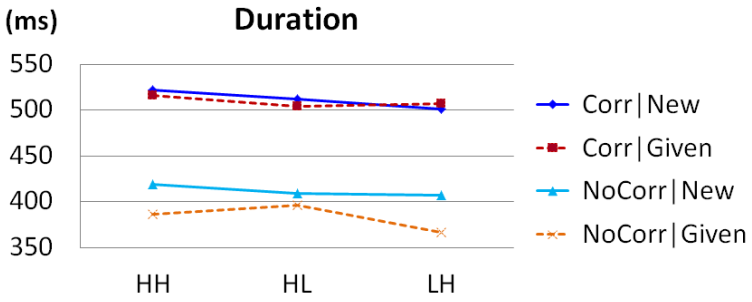
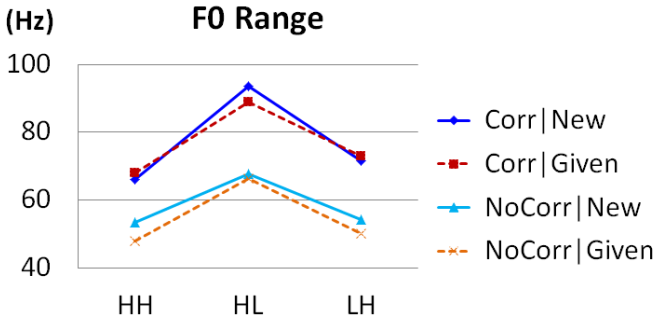
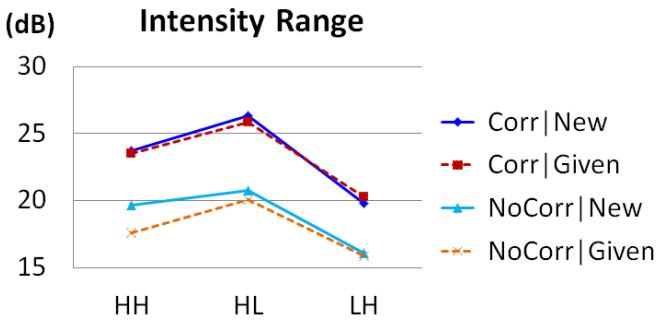


Fig. 2-4: Duration, F0 ranges, and intensity ranges in the four information types (X-axis: tone types of the target words)







References: **[1]** Gussenhoven 1983. Focus, mode, and nucleus. *Journal of Linguistics.* **[2]** Pierrehumbert & Hirschberg 1990. The meaning of intonational contours in the interpretation of discourse. In Cohen et al (Eds), *Intentions in communication.* **[3]** Watson et al 2008. Interpreting Pitch accents in online comprehension. *Cognitive Science.* **[4]** Whalen & Xu 1992. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica.* **[5]** Xu 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics.* **[6]** Greif 2010. Contrastive focus in Mandarin Chinese. *Proceedings of Speech Prosody* 2010. **[7]** Chen & Gussenhoven 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics.*