Focus-marking in a tone language: Prosodic cues in Mandarin Chinese

Iris Ouyang & Elsi Kaiser
Department of Linguistics, University of Southern California, Los Angeles, USA
{ chuoyino, emkaiser } @usc.edu

We conducted a production study to investigate prosodic encoding of discourse information in Mandarin, a tone language where F0, duration, and intensity also distinguish lexical items. We tested whether (i) the presence/absence of correction and (ii) the new/given distinction are encoded prosodically. Our results indicate that correction was reflected in all three parameters: Corrective words had longer durations, larger F0 ranges, and larger intensity ranges than Non-corrective words. The distinction between new and given information was reflected only in duration and F0, and only in the absence of correction (Correction-by-Givenness interaction). We discuss how these findings highlight the ability of the production system to utilize different aspects of acoustic dimensions.

1 Introduction

Prosody conveys discourse-level information (Gussenhoven 1983, Pierrehumbert & Hirschberg 1990) with three main acoustic dimensions, namely duration, intensity, and fundamental frequency (F0). In English, signals of prosodic prominence – such as longer duration, greater intensity, and changes in F0 movement – appear on elements that are semantically or pragmatically prominent in the discourse (Ladd 1996).

Prior work on English has shown that prosody can distinguish between different information structural properties. Katz & Selkirk (2011) show that contrastive focus has stronger effects than new-information focus on words' relative duration, relative intensity, and F0 changes. Concentrating on F0 movement, Watson, Tanenhaus, & Gunlogson (2008) found that an L+H* pitch accent is associated with contrastive focus, whereas an H* accent with both contrastive focus and new-information focus.

Prosodic encoding of information structure becomes more complex when we consider tone languages, where duration, intensity, and F0 also play a role in the lexicon. In Mandarin Chinese, four pitch patterns (referred to as 'tones') work as phonemes: high (Tone 1), rising (Tone 2), low (Tone 3), and falling (Tone 4). They can alter lexical meaning, as shown in (1). In addition to the four-way distinction based on F0 movement, lexical tones in Mandarin also differ in amplitude and length. Tone 2, Tone 3, and Tone 4 are perceptible solely on the basis of their amplitude contours (Whalen & Xu 1992), and Tone 3 is 1.5 times longer than the other tones when produced in isolation (Xu 1997).

(1) Tone 1 ma [High] 'mother'
Tone 2 ma [Rising] 'hemp'
Tone 3 ma [Low] 'horse'
Tone 4 ma [Falling] 'scold'

Prior work suggests that Mandarin resembles English in the way it increases duration and intensity to emphasize information in an utterance, but differs from English in terms of F0 movement (Jin 1996, Chen et al 2009). Instead of imposing pitch accents and changing the *shapes* of F0, narrow focus in Mandarin expands the *ranges* of F0, so that pitch patterns specified for different lexical tones remain distinct within a type of focus (Chen & Gussenhoven 2008). This is presumably due to the fact that the shapes of F0 are the major cue for lexical tones.

The picture is less clear, however, as to (i) whether prosodic cues distinguish different information structural statuses from one another in Mandarin, and (ii) which acoustic dimensions are employed to mark a particular information structural status. For example, does new-information focus differ from contrastive focus? If so, do all three acoustic dimensions provide cues for this distinction? Existing work in this

area has led to somewhat divergent results (*e.g.* Jin (1996) and Chen et al. (2009) on mean intensity), and researchers have also defined focus in different ways (*e.g.* Chen & Gussenhoven (2008) and Grief (2010) on corrective focus).

2 Aims of this study

In this paper, we report a psycholinguistic experiment that aimed to answer two questions: First, *how is information structure encoded prosodically in a tone language*, where all the three acoustic dimensions – duration, F0, and intensity – already serve lexical purposes? Existing work is mostly devoted to duration and F0 (Xu 1999, Chen 2006, Chen & Gussenhoven 2008, Greif 2010). Studies which presented results on intensity only looked at mean intensity (Jin 1996 and Chen et al. 2009). Given that intensity contours, as well as F0 contours, are associated with lexical tones, we wanted to investigate whether intensity ranges can mark discourse-level information just like F0 ranges do. Second, *are different kinds of information structure distinguished prosodically?* Existing work that examined multiple types of information structure in Mandarin mostly concentrated on subtypes of contrastive focus (Chen & Gussenhoven 2008, Greif 2010). The two major types of information structure that have received crosslinguistic attention (*e.g.* Pierrehumbert & Hirschberg 1990, Watson et al. 2008, Katz & Selkirk 2011) – new information and contrastive focus – have not been carefully compared in Mandarin. Building on prior work, we examined the acoustic correlates of (i) new-information focus and (ii) a type of contrastive focus, namely correction focus. We also investigated how the given-new distinction interacts with the presence/absence of corrective focus. Duration, intensity ranges, and F0 ranges were analyzed.

3 Method

In a production study, eight native speakers of Beijing Mandarin, four women and four men, produced instructions to move an object to a location, based on pictures and arrows on the computer screen. For example, in Figure 1, the arrow points from the bamboo to the fridge, so participants should say: 'Move the bamboo next to the fridge.' After they produced the instruction, participants saw a moving event that correctly or wrongly responded to their instruction. They were told to check if their instructions were carried out correctly, and to provide a correction if their instructions were not followed.

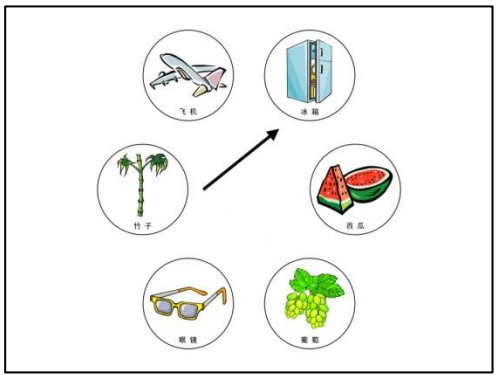


Figure 1: Sample display

To investigate focus-driven intonation across lexical tones, we manipulated the information structure of a target word and controlled its tonal combination. A repeated-measures design with two independent variables was used: (i) **correctiveness** (presence/absence of correction) and (ii) **givenness** (new/given information). Target words were bisyllabic, with High-High (HH), High-Low (HL), or Low-High (LH) tonal contours. All sentences were produced in the frame illustrated in (2)[1]. A target word always appeared in the OBJECT role in a sentence. There were 36 target trials and 36 filler trials.

(2) ba OBJECT fang-dao/-zai LOCATION pangbian
BA OBJECT put-PREP LOCATION side
'Move the OBJECT next to the LOCATION'

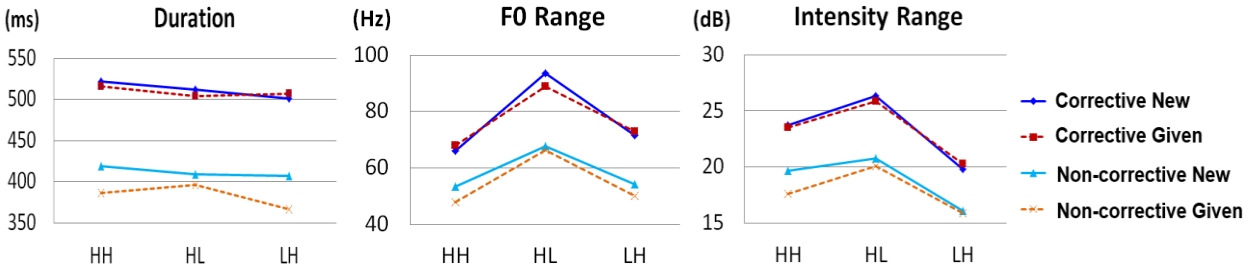
[1] For instance, the sentence "*ba zhuzi* ('bamboo') *fangdao bingxiang* ('fridge') *pangbian*" would be produced for the display in Figure 1. In the verb-preposition construction, *-dao* and *-zai* are interchangeable across speakers in this context. Participants were asked to use the one most natural to them.

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 in order to correct the wrong moving event. In **New** conditions, the target word had not been involved in a moving event on that trial (*i.e.* it was hearer-new). In **Given** conditions, the target word occurred in a correct moving event earlier on the same trial (*i.e.* it was hearer-old). Thus, there were four conditions: Corrective New, Corrective Given, Non-corrective New, Non-corrective Given.

4 Results

Acoustic analyses were done using the Praat software with the ProsodyPro script (Xu 2005-2011). Length, maximum and minimum F0, and intensity were extracted by the script. Repeated measures ANOVAs and paired t-tests were conducted on the duration, intensity ranges (maximum intensity – minimum intensity), and F0 ranges (maximum F0 – minimum F0) of target words.

The **presence and absence of correction** was reflected in all three acoustic dimensions: Target words in the Corrective conditions were *longer*, and had *larger intensity ranges* as well as *larger F0 ranges* than those in the Non-corrective conditions. ANOVAs show main effects of correction on duration, intensity ranges, and F0 ranges (p's< .05). T-tests reveal that these effects of correction occur regardless of whether the target word is new (p's < .05) or given (p's< .01), as can be seen in Figures 2-4.



Figures 2-4: Duration, F0 ranges, and intensity ranges in the four information types (x-axis: lexical tones of the target words)

The **distinction between given and new information** was reflected only in duration and F0, not in intensity, and only *in the absence of correction*. In other words, when the target word was correctively focused, there were no differences between new and given information. However, the distinction between new and given emerged in the Non-corrective conditions, where target words conveying New information were *longer and had larger F0 ranges* than those conveying Given information. As expected based on the patterns visible in Figures 2-4, ANOVAs show that there is no main effect of givenness on duration, intensity ranges, or F0 ranges (p's > .075). Nevertheless, there is an interaction between correction and givenness on duration and F0 ranges (p's< .05), although not on intensity ranges (p= .40). T-tests show that the givenness effects on duration and F0 ranges emerge when the words are non-corrective (p's< .05) but not when the words are corrective (p's> .89). In contrast, no givenness effect is found on intensity ranges in either non-corrective words (p= .31) or corrective words (p= .94).

5 Discussion

This study investigated prosodic cues for two types of information structure in Mandarin Chinese: corrective focus and new-information focus. As we saw in Figures 2-4, correction

yields lengthening, F0 range expansion, and intensity range expansion, whereas new information only triggers lengthening and F0 range expansion, but does not affect intensity range.

In tone languages, discourse-level intonation and lexical tones occupy the same three acoustic dimensions: F0, intensity, and duration. Consistent with prior work, we found lengthening and F0 range expansion in both corrective focus (e.g. Chen 2006, Chen & Gussenhoven 2008) and new-information focus (e.g. Jin 1996, Xu 1999). Furthermore, our results show that intensity ranges may also be expanded to emphasize words in an utterance: Intensity excursions become larger when speakers express a correction. In other words, there is no evidence for specialized functions where some prosodic dimensions mark information structure and others mark lexical items. Instead, *all three prosodic dimensions are multi-functional*.

Our findings about intensity range expansion provide insight into how information structure and lexical items are encoded in prosody. Existing work has pointed out that (a) for different lexical tones, the shapes of F0 contours are clearly distinct, whereas (b) with information structure, what varies are the ranges of F0 contours (Xu 1997, Chen & Gussenhoven 2008). Whalen & Xu (1992) suggest that intensity and F0 are positively correlated in lexical tones. As our results show that intensity ranges are used to convey discourse focus, there appear to be parallels between F0 and intensity in the specialization of parameters. Lexical information is encoded in the *shapes* of F0 and intensity movement, whereas discourse information is marked by the *ranges* of F0 and intensity movement. This highlights the fine-grained ability of the production system to utilize different aspects of acoustic dimensions.

Given that prosodic cues are present for information structure in a tone language, the next question that comes up is: Do prosodic cues distinguish one type of discourse information from another in Mandarin, like the correspondence between pitch accents and focus types in English? Our results suggest that new-information focus is encoded differently from corrective focus: While both correction and new information lengthen the focused words and expand F0 ranges, only correction is associated with intensity range expansion. The distinction between new and given information only emerges in the absence of correction, however, which may result from several possible reasons and deserves further investigation.

References

Chen. 2006. Durational adjustment under corrective focus in Standard Chinese. *Journal of Phonetics*, 34, 176-201.

Chen & Gussenhoven. 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 36, 724-746.

Chen, Wang, & Xu. 2009. Closely related languages, different ways of realizing focus. *Proceedings of Interspeech 2009*. Brighton, UK.

Greif. 2010. Contrastive focus in Mandarin Chinese. In *Proceedings of Speech Prosody* 2010, Chicago.

Gussenhoven. 1983. Focus, mode, and nucleus. *Journal of Linguistics*, 19, 377-417.

Ladd. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Jin. 1996. *An acoustic study of sentence stress in Mandarin Chinese*. Ph.D. dissertation, the Ohio State University.

Katz & Selkirk. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, 87, 771-816.

Pierrehumbert & Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Cohen et al (eds.), *Intentions in communication*, 271-311. Cambridge, MA: MIT Press.

Watson, Tanenhaus, & Gunlogson. 2008. Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science*, 32, 1232-1244.

Whalen & Xu. 1992. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49, 25-47.

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

Xu. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, 27, 55-105.

Xu. 2005-2011. ProsodyPro.praat. Available from: http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/