

# Individual differences in Mandarin focus production

Yike Yang<sup>1</sup>, Si Chen<sup>1,2</sup>

<sup>1</sup>Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR

<sup>2</sup>The Hong Kong Polytechnic University-Peking University Research Centre on Chinese Linguistics, Hong Kong SAR

<https://doi.org/10.36505/ExLing-2020/11/0057/000472>

## Abstract

This paper investigated whether and how individual speakers of Mandarin Chinese (Mandarin) mark prosodic focus (broad focus vs verb focus) differently in their production, and tested focus effects on mean F<sub>0</sub>, duration and intensity. The findings indicated the role of the three acoustic cues in Mandarin focus marking at both the group and individual levels. Meanwhile, the individual data showed great variations among speakers in terms of the extent to which the cues were employed. It is proposed that the dynamics of acoustic cues should be considered in future studies and caution should be taken when selecting stimuli for focus perception studies.

Keywords: individual differences, speech production, prosody, focus, Mandarin

## Introduction

Research into individual speaker differences in speech production is of great significance for our understanding of the way speakers signal linguistic contrasts (Smith & Hawkins 2012). However, little is known about how individual speakers vary in using prosodic cues to mark information structure (Ouyang & Kaiser 2015). Cross-linguistically, various acoustic cues are employed to signal information structure (Lee 2015). In Mandarin Chinese (Mandarin), focus influences the mean F<sub>0</sub>, duration and intensity in both local and global regions, wherein the on-focus region receives the greatest prominence and the post-focus region receives the least (Yang & Chen 2020).

This project aimed to investigate whether and how individual speakers of Mandarin mark prosodic focus differently in their production. Specifically, this study examined the individual differences in the implementation of mean F<sub>0</sub>, duration and intensity in pre-, on- and post-focus regions in the production of Mandarin broad focus and verb focus.

## Methodology

Eleven native speakers of Mandarin (six females; aged:  $24.72 \pm 4.39$ ) attended a production experiment at the Speech and Language Sciences Laboratory of the

Hong Kong Polytechnic University. All participants gave their written informed consent prior to the recording sessions.

Table 1. Focus types and question-answer pairs (focused regions underlined).

Focus types	Precursor questions	Answers
Broad focus	ni shuo shen me? 'What did you say?'	na wei yisheng he kafei. 'The doctor drinks coffee.'
Verb focus	na wei yisheng zenme kafei? 'What does the doctor do to the coffee?'	na wei yisheng <u>he</u> kafei. 'The doctor <u>drinks</u> coffee.'

The stimuli consisted of six subject–verb–object (SVO) declarative sentences with exactly the same structure. To make the production data more natural, the broad focus and verb focus utterances were elicited with precursor questions, as illustrated in Table 1. Only the answers were further processed. Relevant acoustic values, including mean F0, duration and intensity of each syllable, were extracted and analysed with linear mixed-effects modelling for each region (pre-, on- and post-focus regions, each corresponding to subject, verb and object positions, respectively).

## Results

The group data are plotted in Figure 1, which revealed very clear focus effects. Linear mixed-effects models suggested that focus significantly increased mean F0 ( $p < .001$ ), duration ( $p < .001$ ) and intensity ( $p = .005$ ) in the on-focus region. There was also a main effect of focus on F0 in the pre-focus region and a main effect of focus on F0 and intensity in the post-focus region ( $ps < .001$ ).

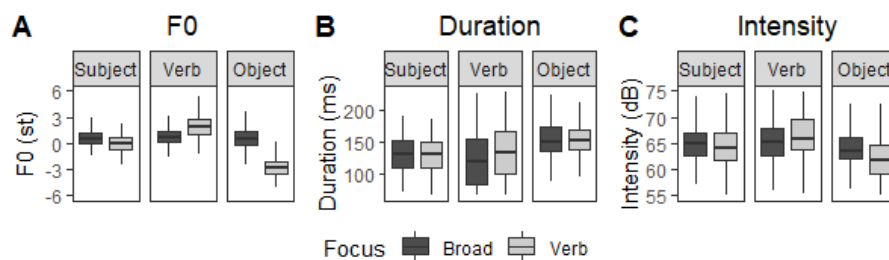


Figure 1. Group data of focus effects.

'Speaker' was then included as a fixed effect to examine whether individual speakers used the acoustic cues differently. Of the nine models (3 regions \* 3 acoustic cues), eight showed a main effect of speaker ( $ps < .001$ ). The only non-significant model was the one for on-focus F0, which reflected the fact that most speakers increased F0 when the constituent was under focus.

Next, individual models were fitted for each acoustic cue produced by each speaker in each region, and the results indicated huge variations among speakers in their use of the acoustic cues, as presented in Figure 2. For example, Speaker 1 did not use F0 ( $p = .715$ ) or intensity ( $p = .890$ ) to mark verb focus in the on-focus region, but there was a clear decrease of F0 and intensity in the pre- and post-focus regions ( $p$ s  $< .039$ ; Figures (2A) and (2C)), which made the pronunciation of the verb prominent. Also, while most speakers increased on-focus F0, the degree of the increase can be as large as  $3.519 \pm 0.396$  st for Speaker 2 ( $p < .001$ ; Figure (2D)) and as small as  $0.445 \pm 0.128$  st for Speaker 10 ( $p = .002$ ; Figure (2G)). Besides, as shown in Figure (2E), Speaker 2 appeared to have lengthened duration in all the regions to mark verb focus.

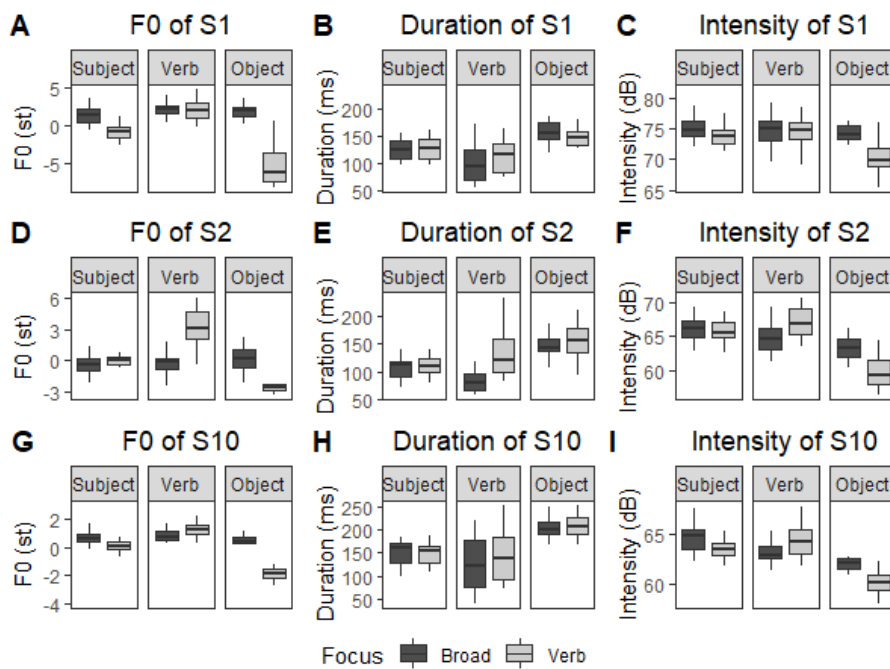


Figure 2. Individual data of focus effects ('S' for 'speaker').

## Discussion and conclusions

This paper investigated how individual speakers differed in marking Mandarin broad focus and verb focus. At the group level, focus affected mean F0 in all the three regions and affected intensity in both on- and post-focus regions, but such effect for duration was observed in the on-focus position only. At the individual level, although every speaker made use of the three tested acoustic cues, there were great variations among speakers in terms of the extent to which the cues were employed.

Note that we converted the F0 values of each speaker from Hz to st individually, with the mean F0 of each speaker as the reference. Despite the fact that normalised scores were used, the F0 values exhibited large variations even under the broad focus condition. If we observe the mean F0 values from the subject position to the object position under broad focus, while Speaker 2 had a gradual increase (Figure (2D)), Speaker 10 showed an increase from subject to verb and then a decrease from verb to object (Figure (2G)). Also, the mean F0 values of Speaker 2 centred around 0 st (Figure (2D)) and the mean F0 values of Speakers 1 and 10 were well above 0 st (Figures (2A) and (2G)). The verb focus condition, needless to say, presented even more complex patterns, which require further examination. In addition to mean values, it is necessary to consider the dynamics of F0 and other acoustic cues to account for the observed individual differences in Mandarin prosodic focus (Roessig et al 2019).

Although the three acoustic cues were used by all the speakers, not every speaker employed these cues in each syntactic position. From our preliminary data, we postulate a possible hierarchy of the cues in marking Mandarin focus: F0 > intensity > duration. There still remain some issues for focus perception studies. First, given the different weights the cues carry in focus production, a follow-up question would be how much each acoustic cue contributes to Mandarin focus perception. Second, having observed very complicated patterns in focus production, one would ask whether there are also interactions among the cues in focus perception. Third, if speaker variation is the norm in focus production, we must be more cautious when selecting stimuli for focus perception tasks.

In conclusion, the findings indicated the role of mean F0, duration and intensity in Mandarin focus marking at both the group and individual levels. The dynamics of acoustic cues should be considered in future studies to account for the observed differences. Besides, issues for focus perception studies are also proposed.

## References

- Lee, Y.-C. 2015. Prosodic focus within and across languages. Unpublished PhD dissertation, University of Pennsylvania.
- Ouyang, I. C., Kaiser, E. 2015. Individual differences in the prosodic encoding of informativity. In Fuchs, S. et al (eds.), *Individual Differences in Speech Production and Perception*, 147-188. Bern, Peter Lang.
- Roessig, S., Mücke, D., Grice, M. 2019. The dynamics of intonation: Categorical and continuous variation in an attractor-based model. *PLoS ONE*, 14(5), e0216859.
- Smith, R., Hawkins, S. 2012. Production and perception of speaker-specific phonetic detail at word boundaries. *Journal of Phonetics*, 40(2), 213-233.
- Yang, Y., Chen, S. 2020. Revisiting focus production in Mandarin Chinese: Some preliminary findings. In *Proc. Speech Prosody 2020*, 260-264, Tokyo, Japan.