Tonal Aspects of Languages 2016
24-27 May 2016, Buffalo, New York

Prosodic focus marking in Dali Mandarin

Zenghui Liu[1], *Hans Van de Velde* [2], *Aoju Chen*[1]

[1] Utrecht University
[2] Fryske Akademy
l.z.h.liu@uu.nl, aoju.chen@uu.nl, HVandeVelde@fryske-akademy.nl

Abstract

This study investigated prosodic marking of focus in Dali Mandarin, a variety of Xinan Guanhua (Southwestern Mandarin) spoken in Dali city, the capital of Dali Bai Autonomous Prefecture, China. Dali Mandarin as a variety of Mandarin has had heavy contact with Bai, a Tibeto-Burman language, for a long time. We adopted a semi-spontaneous experimental approach to elicit SVO sentences with different focus conditions. Our data showed that native speakers of Dali Mandarin lengthened the duration of focal constituents compared to non-focal constituents for marking focus. However, they did not use duration to distinguish focus types differing in size and contrastivity. Further, pitch played no role at all in signaling focus, or in differentiating focus types. These results thus suggested that Dali Mandarin speakers use prosody by exploiting duration to mark focus. Therefore, the encoding of focus in Dali Mandarin is more similar to Bai than to Beijing Mandarin, the latter being genetically more closely related to Dali Mandarin. This result suggests that prosodic focus marking in Dali Mandarin has been influenced by Bai due to the heavy language contact.

Index Terms: focus, intonation, language contact, Dali Mandarin

1. Introduction

Focus refers to the new information in a sentence delivered from a speaker to a listener [1, 2]. Focus can be realized by using different linguistic strategies, such as prosody. Prosodic cues for focus, such as pitch and duration, are language specific [e.g., 3, 4, and 5]. It has also been shown that the use of pitch or duration is not related to the tonal aspects of a language [6]. For instance, both tone (e.g., Mandarin [4, 7], Vietnamese [8]) and non-tone languages (e.g., English [3], Dutch [5], German [9]) use pitch range and duration for focus-marking purposes, whereas other languages only use duration to mark focus, such as Cantonese [10], Yi [11], Tsat [12], Bai [13] and Deang [11]. In addition, there are tone languages that do not use prosodic cues (pitch or duration) at all to mark focus, e.g., Yucatec Maya [14, 15].

Recently, researchers have shown an increased interest in investigating the prosodic realization of focus in dialectal varieties of a language. For instance, a number of studies examined the similarities and differences of prosodic focus marking in different varieties of Mandarin, including the varieties spoken in Nanchang (Gan) [11], Lan Zhou [16], Jinan, Liaocheng, Zibo, Dalian, Harbin, Tianjin and Xi'an [11, 16, 17, 18]. In general, these investigations showed a similar manner of prosodic focus marking in these varieties and Beijing Mandarin, the standard variety of Mandarin. Namely, speakers encoded focus either by raising the mean pitch of the focal constituent or by expanding the pitch range relative to the non-focal constituent. In addition, the compression of the pitch range of the post-focal constituent has also been consistently found in all these varieties of Mandarin.

However, limited research on varieties of Mandarin spoken in the vicinity of a language other than Mandarin have revealed different uses of prosody in focus marking compared to Beijing Mandarin. For example, [19] investigated focus marking in Taiwan Mandarin, compared to Beijing Mandarin and Taiwanese. Taiwan Mandarin, spoken in Taiwan, has been in close contact with Taiwanese for several decades. [19] showed that Taiwan Mandarin is more similar to Taiwanese than to Beijing Mandarin in terms of using prosodic cues for marking focus. Specifically, Taiwan Mandarin monolinguals not only expanded pitch range, but also increased the intensity and duration of the focal constituents for marking focus as Beijing Mandarin and Taiwanese speakers did. However, neither Taiwan Mandarin speakers nor Taiwanese speakers produced the post-focal constituents with compressed pitch range and intensity, which was present in the Beijing Mandarin speakers' production. It was suggested that these differences between Taiwan Mandarin and Beijing Mandarin can be attributed to Taiwan Mandarin's close contact with Taiwanese.

In order to shed more light on the impact of language contact on prosodic focus marking, we investigate how speakers of Dali Mandarin mark focus prosodically. Dali Mandarin is a variety of Xinan Guanhua (Southwestern Mandarin) spoken in Dali city, the capital of Dali Bai Autonomous Prefecture, China. Dali Mandarin has been in heavy contact with Bai for centuries, a Tibeto-Burman language containing eight lexical tones. Both Bai and Dali Mandarin are commonly used by Bai and Han people in Dali.

It has been shown that Bai only uses duration, which is different from Standard Mandarin (Beijing Mandarin) and other varieties of Mandarin that use both pitch and duration to mark focus [13]. Until now the prosodic focus marking system of Dali Mandarin has not been studied. The present study investigates for the first time the prosodic focus marking in Dali Mandarin. Specifically, we examine how pitch and duration may be used to mark focus in Dali Mandarin by finding out (1) whether native speakers of Dali Mandarin mark focus prosodically by varying pitch and duration, i.e., the effect of focus; (2) whether they use prosodic cues to distinguish focus type that differs in the size of focal constituent, i.e., the effect of size; and (3) whether they distinguish contrastive focus from non-contrastive focus using prosodic cues, i.e., the effect of contrastivity.

2. Methodology

2.1. Experimental materials

A semi-spontaneous approach was adopted from [5] to elicit focus marking in SVO sentences. There were five different focus conditions: narrow focus on the subject NP in sentence-initial position: NF-i, example (1); narrow focus on the verb in sentence-medial position: NF-m, example (2); narrow focus on the object NP in sentence-final position: NF-f, example (3); broad focus: BF, example (4); and contrastive focus on the

http://dx.doi.org/10.21437/TAL.2016-22

verb in sentence-medial position: CF-m, example (5). The focus conditions were elicited by a WH-question or a statement from the experimenter, as illustrated in examples (1) to (5), where focused constituents appear in square brackets.

(1) Experimenter: Look! The ball. There is also a waving arm. It looks like someone throws the ball. Who throws the ball?
Participant: [THE RABBIT] throws the ball. (NF-i)
(2) Experimenter: Look! The rabbit and the ball. It looks like that the rabbit does something with the ball. What does the rabbit do with the ball?
Participant: The rabbit [THROWS] the ball. (NF-m)
(3) Experimenter: Look! The rabbit, it waves its arm. It looks like that the rabbit throws something. What does the rabbit throw?
Participant: The rabbit throws [THE BALL]. (NF-f)
(4) Experimenter: Look! This picture is very blurry. I can't see anything clearly. What does the picture show?
Participant: [THE RABBIT THROWS THE BALL]. (BF)
(5) Experimenter: Look! The rabbit and the ball. It looks like that the rabbit does something with the ball. I guess the rabbit cuts the ball.
Participant: The rabbit [THROWS] the ball. (CF-m)

The sound system of Dali Mandarin is similar to Beijing Mandarin [21, 22]. Dali Mandarin has four citation tones, including a mid-high level tone (T1/44), a mid-falling tone (T2/31), a high-falling tone (T3/53) and a low dipping tone (T4/213) [22]. Stimuli included all four lexical tones (Table 1), and the target verbs are in bold.

Table 1. *The composition of the stimuli, split up by tone and syntactic category*

	Tone			
	T1/44	T2/31	T3 /53	T4/213
Subject	猫 cat	熊 bear	狗 dog	兔 rabbit
Verb	**扔** **throw**	**埋** **bury**	**剪** **cut**	**运** **transport**
Object	书 book	球 ball	笔 pen	菜 vegetable

All four tones listed in the Table 1 were systematically balanced in subject noun phrase, the verb and the object noun phrase. Sentence-medial verbs were treated as our targets for acoustic analysis, as the multiple roles played by verbs in various focus conditions. Specifically, the verb was the focal constituent, when the sentence focus condition was NF-m; and the verb was post-focal and pre-focal constituent respectively while in NF-i and NF-f conditions. In total, 80 sentences, 16 for each focus condition, were elicited from each participant.

2.2. Data elicitation

As this is a part of a research that studies the acquisition of prosodic focus marking in Mandarin by Bai children, a picture matching game paradigm was adopted to ensure comparability between children and adults. In the picture matching game, three piles of pictures were used: the experimenter and the participant each held a pile of pictures sorted in a certain sequence; the third pile of pictures was in random order. In the experimenter's pictures (the first pile), either a subject, an action (verb) or an object was missing. The participant's pictures (the second pile) all contained a complete event. The participant's task was to help the experimenter to find the missing part and the experimenter then paired a picture from the third pile to her own following the production of the participant. The participants were explicitly instructed (1) to respond with full sentences and (2) not to show their own pictures to the experimenter. Five practice trials were given before the experiment to familiarize the participants with the game.

In order to ensure the consistency in the participants' word choice, the picture-matching game was preceded by a picture-naming task, which aimed at familiarizing the participants with the target words and the entities in the pictures used in the game.

2.3. Participants and procedure

Six native speakers of Dali Mandarin (three male and three female, aged between 28 and 54) took part in the experiment. The participants all (1) acquired Dali Mandarin as their native language; (2) were using Dali Mandarin on a daily basis with self-estimated daily use exceeding 60%; (3) had no self-reported speech and hearing impairments. All the participants acquired Putonghua (Standard Mandarin) as their second language at the age of 6, and they all identified themselves as native speakers of Dali Mandarin.

The participants were tested individually in a quiet room either in the experimenter's apartment or in the participant's home. The experimenter was a female native speaker of Dali Mandarin (27-year-old). The experiments were recorded using a portable ZOOM H1 digital recorder at a 44.1 kHz sampling rate and 16 bit accuracy. Each session was also video-taped. The experiment lasted 20 to 25 minutes per participant.

3. Analysis and Results

3.1. Acoustic analysis

A sentence was included for further analysis only if it contained no self-correction or hesitation and was uttered as a response to the target question. In total 92% of the obtained responses (N = 440) were included in the analysis. These utterances were annotated in Praat [23] and segmented at word level. The landmarks for the onset and offset of verbs were determined with the help of the waveform and spectrogram.

The maximum and minimum pitch were labelled taking the tonal targets into consideration, following [24]. The pitch values of the pitch landmarks and the time values of the word boundaries were subsequently extracted via Praat scripts. Two measures from these values were calculated: word duration (i.e. offset time minus onset time) and pitch range (i.e. maximum pitch minus minimum pitch). 23 responses had to be excluded from the analysis of pitch range, pitch maximum and minimum due to an unreliable measurement of the pitch value.

3.2. Statistical analysis

Statistical analyses were conducted using mixed-effect modeling in R [25, 26]. We were interested in how well *Focus Condition* and *Tone Verb* could explain the variation in word duration, pitch range, pitch maximum and pitch minimum of the target verbs, i.e., the phonetic measurements of the target verbs.

In all models, *Focus Condition* and *Tone Verb* were included as fixed factors, while *Speaker* and *Sentence* were included as random factors. *Focus Condition* always contained two levels in every comparison listed above to answer specific research questions, and it was defined differently in different

comparisons as a fixed factor; and *Tone Verb* had four levels which referred to the four lexical tones of the target verbs. The experimental design contained one item for each lexical tone, i.e., only one word containing one specific lexical tone was included. Dependent variables were word duration, pitch range, pitch maximum and pitch minimum of the target verbs.

When building the models, only factors that significantly improved the previous model were included in subsequent models. The improvement of the model fit was assessed by the difference in -2LL (log likelihood), i.e., a statistically significant difference between these two models was an indication of a significant effect of the added fixed factor. We excluded the models that did not lead to a significant improvement over the previous model to get the best fit model. Using this procedure we could assess the effect of the factors listed, as well as their interactions.

To find out the effect of focus, we compared the duration and pitch-related measurements of the focal constituent to the non-focal constituent, i.e. NF-m (focus) vs. NF-i (post-focus); NF-m (focus) vs. NF-f (pre-focus). The effect of size was studied by comparing the duration and pitch-related measurements of the verbs in the narrow focus condition (NF-m) to their counterparts in the broad focus condition (BF). The effect of contrastivity was operationalized by comparing contrastive (CF-m) to non-contrastive focus (NF-m).

3.3. Results

3.3.1. Duration

Duration wise, the verbs were on average 13.8 ms longer when focused (NF-m, M = 213.7 ms, sd = 46) than when not focused and following a focused constituent (NF-i, M = 199.9 ms, sd = 32.1). Mixed-effect modeling was used to assess the effect of *Focus Condition* on the duration of the verbs, as described above. It revealed a main effect of *Focus Condition* ($\chi^2(1) = 5.9477$, $p < .05$). The best fit model contained main effects of *Focus Condition* and *Tone Verb.* This suggested that the duration of the verbs in the focus position were lengthened in all the lexical tones in comparison to their counterparts in the post-focus position.

The use of duration for distinguishing the verbs in the focus position from the post-focus position is showed in Figure 1.

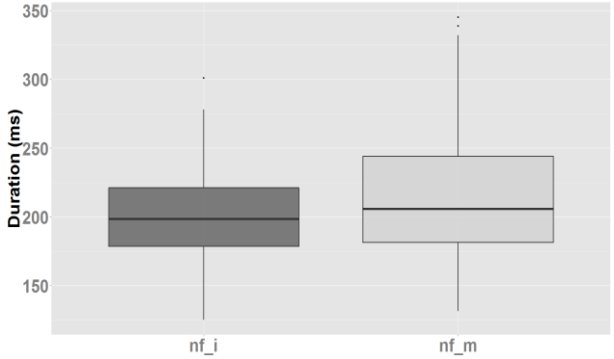

Figure 1: *Mean duration (in ms) of verbs in post-focus vs. focus position. (nf-i = verb following a focused constituent, nf-m = verb in sentence-medial focused position)*

The duration data obtained from the verbs in the NF-m and NF-f conditions showed that the verbs were on average 15.2 ms longer when focused (NF-m) than when not focused and preceding a focused constituent (NF-f, M = 198.5 ms, sd = 34.3). Mixed-effect modeling revealed a main effect of *Focus Condition* ($\chi^2(1) = 8.3277$, $p < .01$). Neither the main effect of *Tone Verb* ($p = .15$), nor the interaction between *Focus Condition* and *Tone Verb* ($p = .53$) significantly improved the model fitness. This suggested that the duration of the verbs in the focus position were lengthened in all the lexical tones in comparison to their counterparts in the pre-focus position.

The use of duration for distinguishing the verbs in the focus position from the pre-focus position is showed in Figure 2.

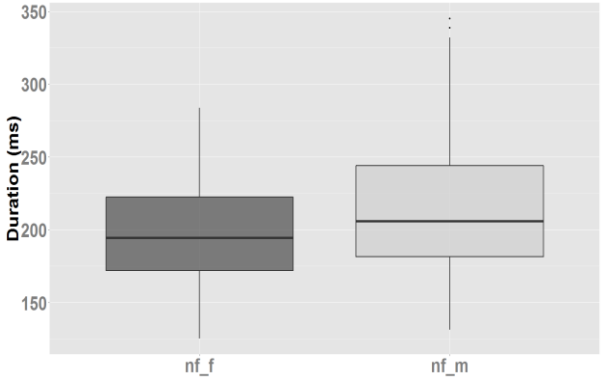

Figure 2: *Mean duration (in ms) of verbs in pre-focus vs. focus position. (nf-f = verb preceding a focused constituent, nf-m = verb in sentence-medial focused position)*

With regard to the effect of focus type that differs in size, we compared narrow focus (NF-m) to broad focus (BF). Mixed-effect modeling revealed no main effect of *Focus Condition* ($p = .08$). Further, there was no interaction between *Focus Condition* and *Tone Verb* ($p = .64$). Thus, duration was not used to differentiate focus type that differs in size.

With regard to the effect of focus type which differs in the constrastivity, we compared contrastive focus (CF-m) to non-contrastive focus (NF-m). Mixed-effect modeling still did not reveal either a main effect of *Focus Condition* ($p = .17$), or an interaction between *Focus Condition* and *Tone Verb* ($p = .86$).

3.3.2. Pitch-related measurements

Mixed-effect modeling confirmed that none of the pitch-related measurements was used in any way for marking focus in Dali Mandarin.

4. Conclusion and Discussion

The present study examined the prosodic focus marking in Dali Mandarin, a variety of Mandarin spoken in the Bai area. Bai and Dali Mandarin have been in contact for centuries. Given that Bai is different from Beijing Mandarin in terms of realizing focus prosodically, the investigation of prosodic focus marking in Dali Mandarin provides us with an opportunity to understand the impact of language contact on prosodic focus marking. As previous studies on Beijing Mandarin [7] and Bai [13] have adopted the same experimental methodology with the present study, the results from these studies are more comparable and are summarized in Table 2.

Table 2. *An overview of prosodic focus marking in BJ Mandarin (Beijing Mandarin), Dali Mandarin and Bai*

Languages	Prosodic cues	Effect of		
		Focus	Size	Contrastivity
BJ Mandarin	Pitch	√	×	×
	Duration	√	√	×
Dali Mandarin	**Pitch**	×	×	×
	Duration	√	×	×
Bai	**Pitch**	×	×	×
	Duration	√	×	×

Given that Beijing Mandarin [4,7] and other varieties of Mandarin [11,16,17,18] exploit pitch and duration as prosodic cues for marking focus, Dali Mandarin as a variety of Mandarin seems to be quite different from other varieties. The prosodic realization of focus in Dali Mandarin is more similar to Bai. Both Bai and Dali Mandarin exploit duration as the major prosodic cue for encoding focus. Thus, our results are in line with previous findings concerning prosodic focus marking in Taiwan Mandarin [19], which is more similar to Taiwanese than to Beijing Mandarin, although Taiwan Mandarin is genetically more closely related to Beijing Mandarin than Taiwanese.

Taken together, these results suggest that language contact can have a significant impact on the prosodic manifestation of focus in languages. Our results show that Dali Mandarin has been influenced by Bai in terms of prosodic focus marking, although the influence of Mandarin on Bai is more well-known. It shows the bidirectional impact on languages which are in the language contact situation.

5. Acknowledgements

We are grateful to all the participants for their cooperation. We also thank Yunshou Liu and Yuhe Yao for their enormous support for the field trip, Anqi Yang and Anna Sara H. Romøren, for their feedback. This study is supported by a scholarship from the Chinese Scholarship Council to the first author and a VIDI grant (276-89-001) from the Netherlands Organization for Scientific Research to the third author.

6. References

[1] Lambrecht, K., Information structure and sentence form: Topic, focus, and the mental representations of discourse referents, vol. 71. Cambridge university press, 1996.

[2] Vallduví, E., and Engdahl, E., "The linguistic realization of information packaging," Linguistics, vol. 34, no. 3, pp. 459–520, 1996.

[3] Cooper, W. E., Eady, S. J., and Mueller, P. R., "Acoustical aspects of contrastive stress in question–answer contexts," The Journal of the Acoustical Society of America, vol. 77, no. 6, pp. 2142–2156, 1985.

[4] Xu, Y., "Effects of tone and focus on the formation and alignment of f0contours," Journal of Phonetics, vol. 27, no. 1, pp. 55–105, Jan. 1999.

[5] Chen, A., "The phonetics of sentence-initial topic and focus in adult and child Dutch," Phonetics and Phonology: Interactions and interrelations, vol. 306, p. 91, 2009.

[6] Chen, S. W., Wang, B., and Xu, Y., "Closely related languages, different ways of realizing focus.," in Interspeech, pp. 1007–1010, 2009.

[7] Yang, A., and Chen, A., "Prosodic focus marking in child and adult Mandarin Chinese," in the 4th International Symposium on Tonal Aspects of Language, pp. 54–58, 2014.

[8] Jannedy, S., "Prosodic focus in Vietnamese," Interdisciplinary studies on information structure, vol. 8, no. 2007, pp. 209–230, 2007.

[9] Féry, C., and Kügler, F., "Pitch accent scaling on given, new and focused constituents in German," Journal of Phonetics, vol. 36, no. 4, pp. 680–703, 2008.

[10] Wu, W. L., and Xu, Y., "Prosodic focus in Hong Kong Cantonese without post-focus compression," Speech Prosody 2010, pp. 1–4, 2010.

[11] Wang, B., Wang, L., and Qadir, T., "Prosodic realization of focus in six languages/dialects in China," Proc. 17th ICPhS, Hong Kong, China, pp. 144–147, 2011.

[12] Wang, B., Li, C., Wu, Q., Zhang, X., Wang, B., and Xu, Y., "Production and Perception of Focus in PFC and non-PFC Languages: Comparing Beijing Mandarin and Hainan Tsat.," in INTERSPEECH, pp. 663–666, 2012.

[13] Liu, Z., Chen, A. and Van de Velde, H., "Prosodic focus marking in Bai," in Speech prosody, vol. 7, pp. 628–63, 2014.

[14] Kügler, F. and Skopeteas, S., "On the universality of prosodic reflexes of contrast: The case of Yucatec Maya," in Proceedings of the 16th International Congress of Phonetic Sciences, pp. 1025–1028, 2007.

[15] Gussenhoven, C. and Teeuw, R., "A moraic and a syllabic H-tone in Yucatec Maya," Fonología instrumental: Patrones fónicos y variación, pp. 49–71, 2008.

[16] Shen, C., and Xu, Y., "Prosodic focus with post-focus compression in Lan-Yin Mandarin".

[17] 段文君, 贾媛, & 冉启斌. (2013). 山东方言焦点语音实现的共性和差异性特征——以济南, 聊城, 淄博方言为例. 清华大学学报: 自然科学版, (6), 835-838.
(Duan, W., Jia, Y. and Q, Ran., "Internal similarities and differences of phonetic realization of the focus in the Shandong dialect – A case study of the Jinan, Liaocheng and Zibo dialects," Journal of Tsinghua University, Sic & Tech, Vol. 53 (6), 835-838. 2013.

[18] Duan, W., and Jia, Y., "The typology of focus realization of Northern Mandarin," in Chinese Spoken Language Processing (ISCSLP), 2014 9th International Symposium on, pp. 492–496, 2014.

[19] Chen, S., Wang, B., and Xu, Y. "Closely related languages, different ways of realizing focus.," in Interspeech, pp. 1007–1010, 2009.

[20] Chen, Y., Xu, Y., and Guion-Anderson, S., "Prosodic realization of focus in bilingual production of Southern Min and Mandarin," Phonetica, vol. 71, no. 4, pp. 249–270, 2014.

[21] 李琳.(2009). 小议云南下关方言和普通话之间的语音差别. 云南农业大学学报, 第 3 卷,第 1 期
(Li, L., "A brief analysis on voice differences between Yunnan Xiaguan Dialect and Mandarin," Journal of Yunnan Agricultural Univerisity. Vol. 3, No.1, 2009.)

[22] 吴积才, 张莆 (1988). 大理方言简况及音系. 玉溪师专学报, 06 期, 61-83
(Wu, J., Zhang, F., "An overview of Dali Dialect and its sound system," Journal of Yuxi Teachers College. No. 06. pp.61-83, 1988.)

[23] Boersma, P., "Praat, a system for doing phonetics by computer," Glot International, vol. 5, no. 9–10, pp. 341–347, 2001.

[24] Xu, Y., and Wang, Q. E., "Pitch targets and their realization: Evidence from Mandarin Chinese," Speech communication, vol. 33, no. 4, pp. 319–337, 2001.

[25] R Core Team (2014). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, 2014.

[26] Bates, D., Maechler, M., and Bolker, B.M., and Walker, S., "lme4: Linear mixed-effects models using Eigen and S4," ArXiv e-print; submitted to _Journal of Statistical Software, 2014.