PROSODIC REALIZATION OF FOCUS
IN SIX LANGUAGES/DIALECTS IN CHINA

Bei Wang, Ling Wang & Tursun Qadir

Minzu University of China, Beijing, China
bjwangbei@gmail.cn; wangzhenling2118@163.com; tuersunka@gmail.com

ABSTRACT

In this paper, the prosodic realization of focus in six languages/dialects in China were experimentally investigated, which were *Yi* and *Tibetan* (Tibetan-Burma), *Wa* and *Deang* (Mon-Khmer), *Uygur* (Turkic) and *Nanchang* dialect (Gan), all spoken in China. *Tibetan*, *Uygur* and *Nanchang* showed on-focus F_0 raising and post-focus F_0 compression (PFC) whereas no clear F_0 variation due to focus was found in *Wa*, *Deang* and *Yi*. On-focus duration lengthening was found in all the six languages. The distribution and the origin of PFC are discussed.

Keywords: focus, post-focus lowering

1. INTRODUCTION

Focus highlights part of a sentence against the rest of the sentence as motivated by a particular discourse situation [2, 3, 6, 9]. A focused word typically has higher F_0, longer duration and greater amplitude compared to its unfocused counterpart (e.g., for English: [5, 6, 12]; for Mandarin: [9]). More importantly, the same studies have also reported a sharp F_0 lowering and pitch range compression in post-focus words. However, two recent studies [3, 8] showed that post-focus compression (PFC) is absent in Taiwanese, Taiwan Mandarin and Cantonese. These studies raised the question as to how PFC got into Mandarin in the first place. Based on these findings and an overview of relevant studies in the world's 30 languages, Xu, et al. [11] argued that because PFC does not seem to be closely related to aspects such as tone, lexical stress or syntactic marker of focus, it is not likely to emerge automatically in a language. In addition, because it is easier to be lost than gained during language contact, PFC is unlikely to have spread through contact. As a result, they suggested, PFC may have a single historical origin.

To better understand the origin of PFC, a large scale typological study is required to map out its distribution. The current study is an initial step toward this goal. Six widely distributed languages and dialects in China were chosen, which were *Yi* and *Tibetan* (Tibetan-Burma), *Wa* and *Deang* (Mon-Khmer), *Uygur* (Turkic) and *Nanchang* (Gan). Among them, only *Yi* and *Nanchang* are tonal. *Yi* has four lexical tones, which are 55, 44, 33 and 21. *Nanchang* has seven tones, which are 42, 24, 213, 5, 21, 5 (checked) and 2(checked) and a neutral tone. In *Uygur*, word stress is mostly assigned to the second syllable. *Tibetan* (Guide dialect), *Wa* (Parauk dialect), and *Deang* (Pulei dialect) lack lexical stress. These languages therefore vary in many aspects. Geographically, the six languages are also well separated, except for *Wa* and *Deang* (see Fig. 1). The main aim of this paper is to establish whether PFC is present in each of these languages. The distribution and origin of PFC will then be discussed based on the new findings.

Figure 1: The locations of the six languages (adopted from [4]).

2. METHOD

2.1. Experimental paradigm

The question-answer paradigm similar to [3, 8] was used for *Nanchang*, *Deang* and *Tibetan*. The speakers were asked to read a target sentence as an answer to the pre-recorded question with emphasis on the appropriate word. For *Wa* and *Yi*, a native speaker uttered the questions to the speakers face-to-face. The *Uygur* data were collected by asking speakers to read the target sentences emphasizing the word marked in bold formant, after understanding the background contexts. For all the languages, the experimenters were native speakers.

If they noticed any inappropriate utterances, they had asked the speakers to say the sentences again.

2.2. Reading materials

For all the languages, two sentences were constructed (see below). For *Yi, Deang, Wa* and *Nanchang*, four focus conditions were elicited by *wh*-questions: initial, middle, final and neutral focus. *Uygur* and *Tibetan* are verb-final languages, so that the realization of final focus is not so clear with the last verb usually being a weak element in the sentence. To solve the problem, we added one more condition for these two languages, which was a focus on the second last word. Below are the reading materials with word-to-word English translation.

Deang:
yiməŋ ʝaŋ a m̥ai (Yimeng sell sugarcane.)
yiməŋ dʒə li dɛdi aai (Yimeng buy book for Ai)
Wa:
roŋ vɛʔ nap hu̥ liaŋ (Rong take Nap to graze.)
ʔai-mɤŋ hu̥ pɛh mak-kram daɯʔ pɯm (Aimong to pick capsicum in kaleyard.)
Uygur:
anam øjdɛ lɛʁmɛn ɛtti (my-mom at-home noddle made)
naːdɛm ajgylgɛ χɛnzutʃɛ ygɛtti (Nadem to-Ajgylge Chinese taught)
Yi:
Mulɛ sytçɔ kota sy dzʐ̩ɔ (Mule woods in tree cut)
Dzysa tʂhɯmu kota tʂhɯ tsy (Zisa rice-field in rice plant)
Tibetan:
Wami rama ze.(fox goat kill)
Ami nəmu la lawa ni. (mom sister to cloth buy)
Nanchang:
surən mo fagu (Suren touch mushroom)
ʝaŋdoŋ baŋ tçiugə duo çiaŋbao (Jiangdong help Qiuge pull suitcase)

2.3. Recording procedure

All the speakers were recorded individually either in a quiet room in their local town (*Deang* and *Yi*) or in the speech lab at Minzu University of China *(Wa, Tibetan, Uygur* and *Nanchang)*. The experimental sentences were repeated three (*Yi, Wa* and *Uygur*) or five (*Deang, Tibetan* and *Nanchang*) times with a random order for each repetition. The speakers practiced the sentences before the recording. For the languages without writing system (*Deang, Wa* and *Yi*), the speakers were asked to remember the two target sentences well and answer the questions with exactly the same sentences.

2.4. Participants

Deang: Seven native speakers, four males and three females, aged 18-56, living in Santaisan, Yunnan province.
Wa: Seven speakers, four males and three females, aged 18-24. They were from Lingchang, Yunnan province.
Uygur: Eight speakers, four males and four females, aged 18-35. They came from different areas in Xinjiang. However, they all spoke standard Uygur without any noticeable accent.
Yi: Seven speakers, four males and three females, aged 22-45, from Leipo country, Xichang, Yunnan province.
Tibetan: Eight speakers, five female and three male, aged 19-23, all from Guide country, Xining, Qinghai province.
Nanchang: Six speakers, four males and two females, all from Nanchang, aged 20-24.

All the speakers spoke Mandarin Chinese as their second language.

2.5. Acoustic measurement

The individual target sentences were extracted and saved as separate wav files. The procedures of acoustic analysis were similar to those in [9, 12]. A Praat script [10] was used to take F_0 and duration measurements from the target sentences. To extract continuous F_0 contours, the vocal cycles were first marked by Praat [1] and then hand-checked for errors. Segmentation labels were also added to mark the syllable boundaries. The script [10] then computed the highest and lowest F_0 and the duration of each syllable. The F_0 values were converted from Hz to semitone (st).

3. RESULTS

3.1. F_0

For all six languages, the time-normalized F_0 contours of the two target sentences in the four/five focus conditions are presented in Fig. 2, averaged across all speakers' repetitions.

From Fig. 2 we can see that focus causes obvious F_0 variation in *Uygur, Nanchang* and *Tibetan*, but not much in *Yi, Wa* and *Deang*. Moreover, in *Nanchang* and *Tibetan*, post-focus F_0 goes below the neutral condition, whereas in *Uygur*, F_0 in post-focus parts drops rapidly, however it goes to the same level as its neutral counterpart. With a closer look, we can see that the neutral sentence in *Uygur* has a similar intonation contour as its initial focus counterpart, instead of the final focus which is the case for *Tibetan* and *Nanchang*.

For statistic tests, the values of maximum F_0 of each word under all focus conditions were calculated. For each language (except for *Tibetan*), a three-way repeated measures ANOVA was carried out, with word position, focus condition

and sentence as independent variables. For *Tibetan*, two-way repeated measures ANOVAs with word position and focus condition as two independent variables were carried out for short and long sentences separately, because the focus conditions are not the same in short and long sentences. The results are presented in Table 1. Due to space limitation, only the F and *p* values of the effect of focus are presented.

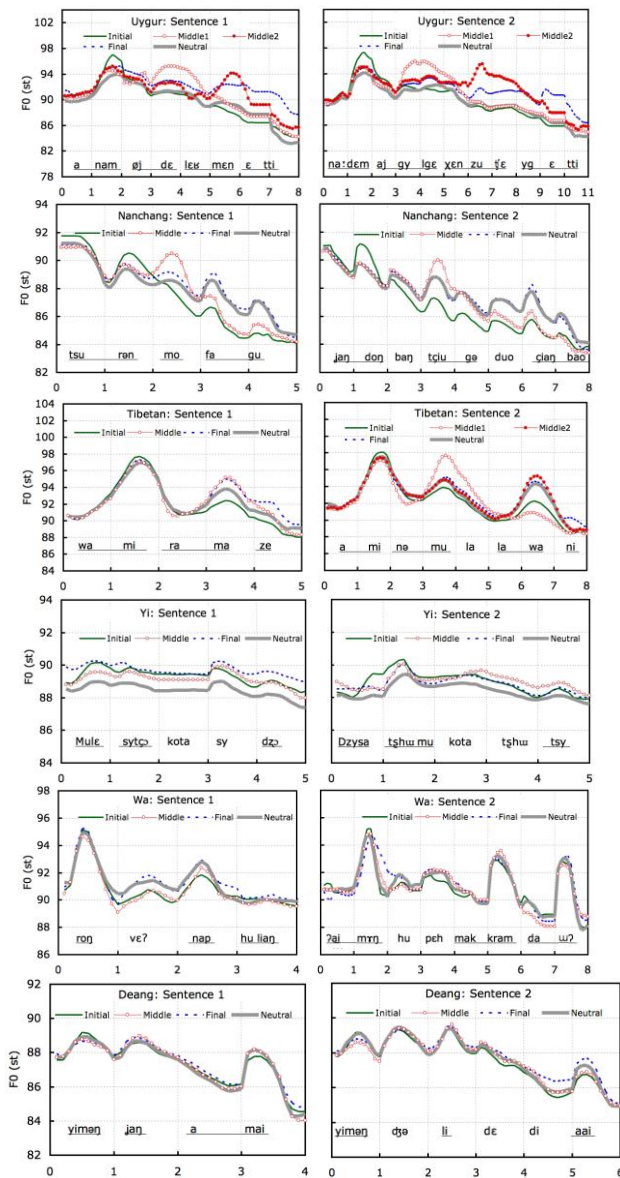**Figure 2:** Intonational contours of the two sentences of the six languages in the four/five focus conditions.

Table 1 shows clearly that focus is encoded with F_0 variation in *Uygur*, *Nanchang* and *Tibetan*, but not in *Yi, Wa* and *Deang*. Although in *Yi*, pitch seems to be raised all through the sentence in focus conditions, it does not reach statistic significance. More importantly, no PFC is shown in *Yi*.

Table 1: The F and p values of the effect of focus by three-way repeated measures ANOVA for maximum F_0 in the six languages/dialects.

	Focus effect
Uygur	$F(4, 28)=13.929$, *p*<.001
Nanchang	$F(3, 15)=15.913$, *p*<.001
Tibetan	Short: $F(3, 21)=8.276$, *p*=.001. Long: $F(4, 28)=14.7$, *p*<.001
Yi	$F(3, 15)=2.742$, *n.s.*
Wa	$F(3, 12)=1.194$, *n.s.*
Deang	$F(3, 18)=1.084$, *n.s.*

Fig. 3 shows averaged pre-focus/on-focus/post-focus pitch change in the six languages/dialects. The pitch change was calculated as the difference of maximum F_0 between pre-focus/on-focus/post-focus and neutral conditions, averaged across the target words and all the speakers. In *Uygur*, no post-focus lowering is found if taken neutral focus condition as the baseline (-0.01 st), because neutral condition is similar to initial focus instead of final focus, so that there is already PFC in neutral condition. It is supported by our perception experiment [7], which shows that the confusion of neutral focus to initial focus is 33.1% but only 6.3% to final focus. For that reason, in Fig. 3, final-focus was taken as the baseline for *Uygur*.

It can be seen in Fig. 3 that PFC shows clearly in *Nanchang, Tibetan and Uygur*, whereas no PFC is found in *Yi, Wa and Deang*. The on-focus F_0 raising is shown in all the languages except for *Deang*.

Figure 3: The difference of maximum F_0 between pre-, on- and post-focus and its neutral counterpart.
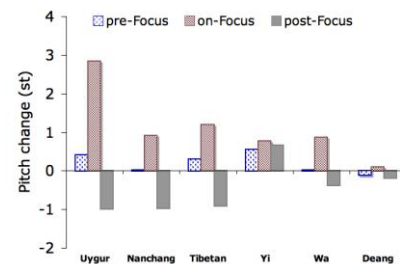
3.2. Duration

Similar to the analysis of maximum F_0, the averaged duration change of words in the pre-focus/on-focus/post-focus conditions relative to the neutral condition is shown in Fig. 4.

The statistic results of duration (similar to that of maximum F_0) are presented in Table 2. As can be seen in Table 2, on-focus duration lengthening is applied in all the languages. In *Nanchang*, *Tibetan*, *Wa* and *Deang*, there is no significant

effect of focus, however the interaction of focus with word position is significant.

Figure 4: The difference of word duration between pre-, on- and post-focus and its neutral counterpart.
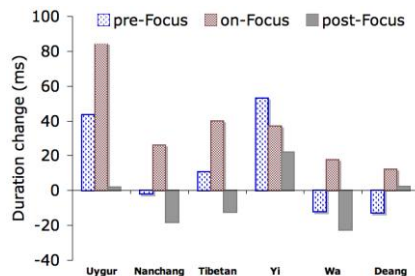

Table 2: The F and *p* values of the effect of focus for word duration in the six languages/dialects.

	Focus effect
Uygur	F(4, 28)=13.558, *p*<.001
Nanchang	W×F: F(6,30)=12.597,*p*<.001
Tibetan	Short:W×F: F(6,42)= 19.336,*p*<.001 Long: W×F: F(12, 84)=12, *p*<.001
Yi	F(3, 15)=5.431, *p*=0.051
Wa	W×F: F(6,42)=3.85,*p*=0.008
Deang	W×F: F(6,36)=5.212,*p*=0.019

Note: W×F stands for the interaction between word position and focus condition.

4. GENERAL DISCUSSION

In line with Taiwanese, Taiwan Mandarin [3] and Cantonese [8], we have found that *Yi, Wa* and *Deang* also lack PFC. *Taiwanese* and *Cantonese* have 6-7 tones, *Yi* has 4 lexical tones, *Wa* and *Deang* are non-tonal. This strengthens the argument by Xu, et al. [11] that the lack of PFC is not related to the tonal aspect of a language.

Another important finding in [3] is that PFC is easy to lose but hard to gain, as Taiwan Mandarin lost PFC due to language contact. In the current study, the speakers of *Yi, Wa* and *Deang* all spoke Mandarin as their second language. But PFC did not transfer from Mandarin to their first languages. We've also tested monolingual speakers of *Deang*, and found that they did not differ from the bilingual speakers. It further supports the argument of Xu, et al. [11] that PFC is not easily spread through language contact.

Wa and *Deang* belong to mon-Khmer language. And, both of them lack of PFC. On the other hand, *Beijing Mandarin*, *Nanchang* and *Tibetan* belong to Sino-Tibetan language family, and they all showed PFC. It indicates that the distribution of PFC may relate to language family.

Xu, et al. [11] proposed that the origin of PFC may come from one proto-language, very likely

being Altaic. And, in this study, we can see that *Uygur* as an Altaic language does have PFC. If that is the case, then PFC can be used as an effective cue for languages classification.

In this study, *Tibetan* shows PFC whereas *Yi* lacks PFC. *Tibetan* and *Yi* are in one language family (Tibeto-Burman). The conflict between language classification and distribution of PFC may lie in the fact that the classification of languages is mostly based on lexical and syntactic features, which can be spread through language contact. However, based on the experimental evidences we had so far, PFC can not be spread through contact.

To further understand the distribution, the origin and the spreading of PFC, a large scale investigation of languages in China is on the way.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] Boersma, P., Weenink, D. 2005. *http://www.fon.hum.uva.nl/praat/*

[2] Bolinger, D.W. 1958. A theory of pitch accent in English. *Word* 14, 109-149.

[3] Chen, S.W., et al. 2009. *Closely Related Languages, Different Ways of Realizing Focus*. Presented at the Interspeech 2009, Brighton, UK.

[4] Chinese languages: *http://www.chinahighlights.com/map/china-city-map.htm*

[5] Cooper, W.E., et al. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America* 77, 2142-2156.

[6] Eady, S.J., Cooper, W.E. 1986. Speech intonation and focus location in matched statements and questions, *Journal of the Acoustical Society of America* 80, 402-415.

[7] Qadir, T., Wang, B. Manuscript. *Production and Perception of Prosodic Focus in Uygur*.

[8] Wu, W.L., Xu, Y. 2010. Prosodic focus in Hong Kong Cantonese without post-focus compression. *Speech Prosody 2010* Chicago, USA.

[9] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *J. of Phonetics* 27, 55-105.

[10] Xu, Y. 2005-2010. TimeNormalizeF0.praat: *http://www.phon.ucl.ac.uk/home/yi/tools.html*

[11] Xu, Y., et al. In press. Prosodic focus with and without Post-Focus Compression (PFC): A typological divide within the same language family? *Linguistic Review*.

[12] Xu, Y., Xu, C.X. 2005. Phonetic realization of focus in English declarative intonation. *J. of Phonetics* 33, 159-197.