

Prosodic Realization of Focus in Statement and Question in Tibetan (Lhasa Dialect)

Xiaxia Zhang¹, Bei Wang^{1*}, Qian Wu¹, Yi Xu²

¹Institute of Chinese Minority Languages and Literatures, Minzu University of China, Beijing, China

²Department of Speech, Hearing and Phonetic Sciences, University College London, London, UK

xiaxia523@163.com, bjwangbei@gmail.com, mdwuqian@qq.com, yi.xu@ucl.ac.uk

Abstract

The present study investigated prosodic realization of focus in statement and yes/no question in Tibetan (Lhasa dialect). Two target sentences were uttered by 8 speakers in both statement and yes/no question under four focus conditions, initial, medial, final and neutral focus. Systematic acoustic analysis showed that: (1) On-focus words exhibited significant F_0 rising, pitch range expansion, and duration lengthening, while pitch range of pre-focus words remained largely intact. Importantly, focus led to post-focus F_0 lowering and compression. Focus was realized the same way in both statement and question. (2) Interrogative intonation had higher F_0 than statement in unfocused words. In general, the prosodic realization of focus in Tibetan was comparable to English and Beijing Mandarin. The distribution and origin of post-focus compression (PFC) was also discussed.

Index Terms: Tibetan, focus, post-focus compression, PFC

1. Introduction

Focus highlights part of a sentence as motivated by a particular discourse situation [1-4]. Cross-linguistic studies have shown that a focused word typically has higher F_0 , longer duration, greater amplitude compared to its unfocused counterpart [2, 5-7]. More importantly, post-focus words exhibit compressed pitch range and intensity. Such post-focus compression (PFC) is present in many languages, including English [2, 6], German [8], Japanese [9], Korean [10], Swedish [11], Beijing Mandarin [5, 12, 13], Uyghur [7] and Tibetan (Ando dialect) [7, 14]. However, PFC is not an universal feature, as it is absent in many African languages [15] and many languages spoken in southern China, e.g. HongKong Cantonese [16], Taiwanese, Taiwan Mandarin [12], Yi [7], Wa [7, 14] and Deang [7, 14]. Such a complex pattern of cross-linguistic distribution of PFC prompted Xu et al. [12, 17] to contemplate a possible historical origin of PFC. Four possibilities are considered: (1) PFC emerged locally in the language — *the local emergence hypothesis*, (2) PFC is correlated to certain linguistic properties, such as lexical tone or stress — *the linguistic co-dependence hypothesis*, (3) PFC is spread across languages through contact — *the spreading hypothesis*, and (4) PFC is inherited from a proto-language — *the inheritance hypothesis*. Of these hypotheses the first one, *local emergence*, is the weakest and so can always be fallen back on when all others

are falsified. The other three all have seen some support but there is no decisive evidence so far. For the *linguistic dependence hypothesis*, the correlation between PFC and lexical tone and lexical stress has already shown to be weak by the fact that both PFC and non-PFC languages have their share of presence as well as absence of lexical tone and stress. For the *spreading hypothesis*, there is already some negative evidence, because PFC is found to be easy to lose but hard to gain when two languages are in close contact. The *inheritance hypothesis* is the most radical one, and so its support would require very strong and comprehensive evidence.

In this study we are interested in providing more experimental evidence on the distribution of PFC through a study of Tibetan, which may in turn contribute to the theoretical discussion on the origin of PFC.

Tibetan belongs to the Sino-Tibetan language family, Tibeto-Burman group. It includes three major dialects, Ando, Lhasa and Kam. Ando is non-tonal, and the other two are tonal [18]. Tibetan is a SOV language. Wang et al. [7, 14] have found that Ando marks focus with a similar tri-zone F_0 adjustment as in Mandarin [5, 12, 13] and English [2, 6]. The exception is that initial focus does not show much F_0 rising.

Interestingly, another Tibeto-Burman language, Yi (also known as Nousu), has been found to lack PFC [7]. Yi is spoken in Sichuan province in the southwest of China. There is no clear F_0 variation due to focus in Yi, but only some on-focus duration lengthening [7]. Ando is mainly spoken in the northwest of China, and Lhasa is spoken in the southwest part, which is close to the Yi speaking area. If PFC is absent in Lhasa, it may suggest that PFC is a geographic area feature instead of a language-specific feature, which would be in support of the *spreading hypothesis* as opposed to the *inheritance hypothesis*. If PFC exists in Lhasa, it would constitute evidence against the *linguistic co-dependence hypothesis* because PFC is not even correlated with the presence of lexical tone across two dialects of the same language.

Another question we are interested in is how different communicative functions are co-encoded in intonation in Lhasa. Liu and Xu have investigated the realization of focus in both statement and question in Mandarin [13] and English [19, 20]. They found that in Mandarin, post-focus pitch range is compressed and lowered in both statement and question. In yes/no questions of English, in contrast, post-focus pitch range is compressed, but raised. Both statement and yes/no question increase their max F_0 and duration in the focused words in these two languages. It would be interesting to find

* Corresponding author

out whether the interaction between focus and question in Lhasa is similar to Mandarin or English.

The present study is to investigate the prosodic realization of focus in statement and question in Lhasa Tibetan. We expect that focus in Lhasa is realized in the same pattern as Ando, but whether there is F_0 increasing on the initial focus is hard to predict.

2. Method

2.1. Materials

Two target sentences were constructed, and each was read as both statement and yes/no question. The only difference between statement and question was that there was an interrogative particle / pɛ / in the final position in question. The lexical tones of Lhasa dialect are still controversial though. We employed Wang's notation [21]. Tones of the target words were also controlled in the two sentences (see (1) and (2)).

(1) $me^{11}to^{251}$ kɛ thɔn¹¹tɕup⁵¹ la thu²¹¹lo²⁵¹ ŋiŋ⁵⁵tɕe¹¹po⁵³ ei²⁵³ nɔ²¹³² pa re (pɛ)

Meto particle Tunzhu particle clothes beautiful bought particles (interrogative particle)

Meto bought beautiful clothes for Tunzhu.

Did Meto buy beautiful clothes for Tunzhu? (with / pɛ /)

(2) $kɛ^{55}saŋ^{51}$ kɛ tɕh⁵⁵riŋ⁵¹ la tɕək⁵⁵ŋu²⁵¹ sa⁵⁵pa⁵³ ei²⁵³ ja⁵⁵ pa re (pɛ)

Kesang particle Ciren particle pen new lent particles (interrogative particle)

Kesang lent a new pen to Ciren.

Did Kesang lend a new pen to Ciren? (with / pɛ /)

There were four focus conditions in the target sentences: initial, medial, final and neutral focus. In total, there were 384 utterances (2 target sentences × 2 sentence types × 4 focus conditions × 3 repetitions × 8 speakers).

2.2. Participants

Eight native speakers of Lhasa dialect, four females and four males, participated in the study. All of them came from Lhasa, except for one male speaker, who was from Sannan, which is next to Lhasa. He spoke standard Lhasa without any noticeable accent. They all spoke Mandarin Chinese as their second language. They were students at Minzu University of China aged between 21 and 24, and did not report any speech or hearing disorders.

2.3. Recording procedure

The recordings were conducted in a sound-proof speech lab, using the AudioRec recording software with Rode NT1-A microphone. All the recordings were directly saved to a Lenovo laptop as wav files. The sampling rate was 44.1 kHz. During the recording, each question and statement pair was displayed on the computer screen, with the focused word being highlighted. The experimenter, a 24-year-old female native Lhasa speaker, had conversations with all the speakers individually. To elicit speakers emphasizing a certain word in statement, the experimenter asked “ $team^{11}jaŋ^{51}$ kɛ thɔn¹¹tɕup⁵¹ la thu²¹¹lo²⁵¹ ŋiŋ⁵⁵tɕe¹¹po⁵³ ei²⁵³ nɔ²¹³² pa re pɛ ? (Did **Kayang** buy beautiful clothes for Tunzhu?)”, and the speaker answered with a corrective focus on the first word:

“ $me^{11}to^{251}$ kɛ thɔn¹¹tɕup⁵¹ la thu²¹¹lo²⁵¹ ŋiŋ⁵⁵tɕe¹¹po⁵³ ei²⁵³ nɔ²¹³² pa re (Meto bought beautiful clothes for Tunzhu)”. To elicit focus in question, the speaker asked with the target sentence “ $me^{11}to^{251}$ kɛ thɔn¹¹tɕup⁵¹ la thu²¹¹lo²⁵¹ ŋiŋ⁵⁵tɕe¹¹po⁵³ ei²⁵³ nɔ²¹³² pa re pɛ? (Did **Meto** buy beautiful clothes for Tunzhu?)”, and then the experimenter answered “ $team^{11}jaŋ^{51}$ kɛ thɔn¹¹tɕup⁵¹ la thu²¹¹lo²⁵¹ ŋiŋ⁵⁵tɕe¹¹po⁵³ ei²⁵³ nɔ²¹³² pa re (**Kayang** bought beautiful clothes for Tunzhu)”. All the sentences were repeated three times in a random order for each repetition and each speaker. There was a short practice session before the recording session.

2.4. Acoustic measurement

The syllabic boundaries of the target sentences were labeled by hand and the vocal cycles were hand-checked. ProsodyPro [22], a Praat script was used to extract the F_0 and duration parameters from the sentences. The F_0 values were converted from Hz to semitone (st) using formula (1) below.

$$f_{st} = 12 \times \log_2(f_0 / 50) \quad (1)$$

3. Results

3.1. Intonation contour

Due to space limitation, only the intonation contours of sentence 1 are presented. The two target sentences show the same pattern.

Fig. 1 and Fig. 2 are the time-normalized F_0 contours of sentence 1 in statement and question with four focus conditions overlaid in one figure. Each value was averaged across 8 speakers and their three repetitions.

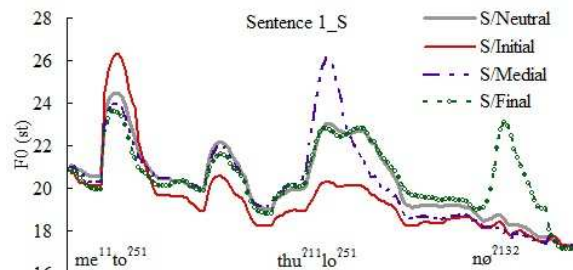


Figure 1: Intonation contour of sentence 1 in statement (S) under different focus conditions.

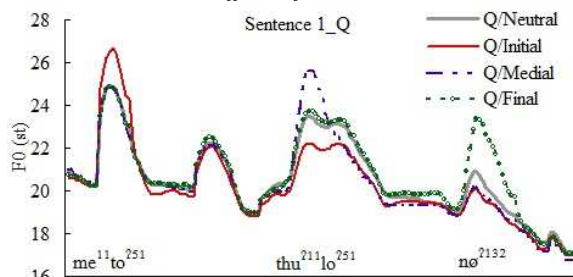


Figure 2: Intonation contour of sentence 1 in question (Q) under different focus conditions.

From Fig. 1 and Fig. 2 we can see that the F_0 of on-focus words is increased, and that the pre-focus words remains largely intact. Besides, post-focus F_0 is compressed and lowered. These observations hold for both statement and question, and in both sentences.

In Fig. 3, the intonation contours of sentence 1 are displayed with statement and question overlaid in four focus conditions respectively. We can see that interrogative intonation is usually higher than the statement, with the largest difference showing in the final target word. With a closer look, we find that the raising of F_0 in question usually lies in the unfocused word. In the focused words, no clear difference can be seen between these two types of sentences. We can also see that, in the neutral focus condition, the difference between statement and question starts from the final target word. In the initial and medial focus conditions, the apparent difference begins immediately after the focused words. While in the final focus condition, such difference starts from the initial word of the sentence, and it mainly shows on the unfocused content words. These observations hold for both sentences.

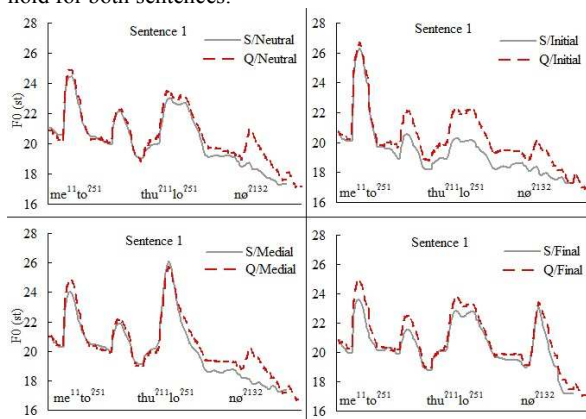


Figure 3: Intonation contours of sentence 1 in statement and question under different focus conditions.

3.2. Quantitative analysis

3.2.1. Max F_0

Table 1 shows maximum F_0 of the three target words in statement and question under different focus conditions, averaged across two sentences.

Table 1. Max F_0 of the three target words in statement and question under different focus conditions (st).

		Word 1	Word 2	Word 3
Statement	Neutral	24.8	23.8	19.9
	Initial	26.7	20.8	19.0
	Medial	24.6	26.2	19.0
	Final	24.3	23.5	23.9
Question	Neutral	25.2	24.0	21.6
	Initial	26.8	22.8	21.1
	Medial	25.2	26.1	20.8
	Final	25.3	24.2	24.0

We can see that focused words have higher maximum F_0 than their unfocused counterpart. Words in post-focus condition have lower maximum F_0 than in the neutral condition. For maximum F_0 of pre-focus words, it differs not much from the neutral counterpart. These are true of both statement and question. Moreover, maximum F_0 of the target words in question is generally higher than that in statement. However, this difference does not show in the focused word.

A three-way repeated measures ANOVA on maximum F_0 , with sentence type, word position and focus condition as

three independent variables. It shows significant main effects of sentence type ($F(1,7) = 22.161, p < 0.05$), word position ($F(2,14) = 33.913, p < 0.001$), and focus condition ($F(3,21) = 17.837, p < 0.05$). In addition, there is an interaction among the three factors ($F(6,42) = 10.465, p < 0.05$).

The follow-up simple effect tests also support the above four observations. In general, focus realization in Lhasa is compatible with that in Mandarin, except that the question/statement contrast in Mandarin is also shown in focused words [13].

3.2.2. Duration

Table 2 displays duration of the three target words under different focus conditions in statement and question, averaged across two sentences.

Table 2. Duration of the three target words in statement and question under different focus conditions (ms).

		Word 1	Word 2	Word 3
Statement	Neutral	317	294	184
	Initial	355	281	182
	Medial	317	327	177
	Final	318	286	227
Question	Neutral	315	286	185
	Initial	346	282	185
	Medial	319	320	179
	Final	313	288	220

As shown in table 2, the three target words have significantly longer duration when they are focused than not focused in both statement and yes/no question.

A similar three-way repeated measures ANOVA on word duration, shows main effects of word position ($F(2,14) = 609.332, p < 0.001$) and focus condition ($F(3,21) = 6.360, p < 0.01$), but not sentence type ($F(1,7) = 0.633, n.s.$). In addition, the interaction between word position and focus condition is also significant ($F(6,42) = 40.403, p < 0.001$).

In summary, in both statement and question, focus increases maximum F_0 and duration of the focused words, and lowers F_0 in the post-focus words, while leaving pre-focus words largely intact. The difference between question and statement mostly lies in F_0 , with interrogative intonation showing higher F_0 in unfocused words.

4. General Discussion

4.1. PFC in Lhasa and its implications

The current data show clear evidence for the presence of PFC in Lhasa Tibetan, just as in Ando Tibetan [7, 14]. This is further evidence that PFC is not related to tonal features of a language, as Ando is non-tonal, and Lhasa is tonal. The falsification of the *linguistic-dependence hypothesis* [12, 17] by this finding is stronger than those made previously, because the tonal diversity happens within the same language in this case.

The present data also show that, although being geographically close to Yi, a non-PFC language, Lhasa has PFC, just like Ando Tibetan. This seems to be at least inconsistent with the *spreading hypothesis*, assuming that geographic affinity correlates with language contact.

Further studies on Tibetan are required to investigate areas with both Tibetan and Yi speakers.

4.2. Co-encoding of focus and interrogative meaning in intonation

From this study, we can see that the realization of focus in statement is very similar in Tibetan (Lhasa and Ando dialect) [7, 14], English [19, 20] and Mandarin [5, 12, 13]. The tri-zone adjustment of focus applies in all these languages, i.e., F_0 is raised in the focused word, and lowered and compressed in the post-focus words, while remaining largely intact in pre-focus words.

Focus realization in yes/no question shows a more complicated pattern among these languages. Lhasa is more similar to Mandarin [13] than to English [19, 20] in the post-focus region, whereas on-focus F_0 rising applies in all these three languages. In English question, post-focus pitch range is compressed but raised. But it is compressed and lowered in Lhasa Tibetan and Beijing Mandarin. The similarity between Mandarin and Lhasa could be due to the fact that they are both tonal. And the underlying mechanism of realizing tone and other communicative functions in intonation is very similar in these two tonal languages.

The difference between Lhasa and Mandarin lies in the adjustment of F_0 rising in interrogative intonation. In Lhasa, the F_0 rising in interrogative intonation mostly show in unfocused words, whereas it starts from the focused words in Mandarin [13]. In English, the F_0 difference between these two sentence types starts from the stressed syllable of the first content word, whether it is focused or not [19, 20].

The current findings in Lhasa is consistent with the PENTA model, which suggests that the individual communicative functions of intonation are encoded in parallel and they are also independent of each other [4].

5. Conclusions

The general finding of the present study is that the encoding of focus in Lhasa is similar to Ando Tibetan, Beijing Mandarin and English, in terms of PFC. However, there are still some systematic differences among them, especially in interrogative intonation with focus. We can draw the following conclusions on the prosodic realization of focus in Lhasa Tibetan:

- (1) The on-focus words exhibit significant F_0 rising, pitch range expansion, and duration lengthening, while the pitch range of pre-focus words remains largely intact. Very importantly, a focus causes post-focus F_0 lowering and compression (PFC). And, it holds for both statement and question.
- (2) Interrogative intonation has higher F_0 than statement in the unfocused words, while there is not much difference between them in the focused words.

6. Acknowledgements

This study was supported by "111" Project of Minzu University. We want to thank Tianlai Data Center for the recording software AudioRec.

7. References

- [1] Bolinger and Dwight, "Intonation and Its Parts," *Palo Alto:Stanford University Press*, 1986.
- [2] S. J. Eady, *et al.*, "Acoustical characterization of sentential focus: narrow vs. broad and single vs dual focus environments," *Language and Speech*, pp. 233-250, 1986.
- [3] R. Ladd, *Intonational phonology*: Cambridge University Press, 1996.
- [4] Y. Xu, "Speech melody as articulatorily implemented communicative functions," *Speech Communication*, vol. 46, pp. 220-251, 2005.
- [5] Y. Xu, "Effects of tone and focus on the formation and alignment of f_0 contours," *Journal of Phonetics*, vol. 27, pp. 55-105, 1999.
- [6] Y. Xu and C. X. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, vol. 33, pp. 159-197, 2005.
- [7] B. Wang, *et al.*, "Prosodic encoding of focus in six languages/dialects in China," presented at the 17th ICPHS, Hong Kong, 2011.
- [8] C. Féry and F. Kügler, "Pitch accent scaling on given, new and focused constituents in German," *Journal of Phonetics*, vol. 36, pp. 680-703, 2008.
- [9] S. Ishihara, "Intonation of Wh-questions in Japanese and its Influence on Syntax," presented at the Linguistics and phonetics 2002, Meikai University, 2002.
- [10] Y.-C. Lee and Y. Xu, "Phonetic Realization of Contrastive Focus in Korean," presented at the Speech Prosody 2010, Chicago, 2010.
- [11] M. Heldner and E. Strangert, "Temporal effects of focus in Swedish," *Journal of Phonetics*, vol. 29, pp. 329-361, 2001.
- [12] Y. Xu, *et al.*, "Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family?," *The Linguistic Review*, vol. 29, pp. 131-147, 2012.
- [13] F. Liu and Y. Xu, "Parallel encoding of focus and interrogative meaning in Mandarin intonation," *Phonetica*, vol. 62, pp. 70-87, 2005.
- [14] Wang, L. "Prosodic marking of focus: Comparing Deang, Wa, Tibetan and Mandarin," Master thesis. Minzu University of China. 2011. (in Chinese)
- [15] S. Zerbian, "Investigating prosodic focus marking in Northern Sotho," in *Focus strategies: evidence from African languages*, K. Hartmann and M. Zimmermann, Eds., ed Berlin: Mouton de Gruyter, 2007, pp. 55-79.
- [16] W. L. Wu and Y. Xu, "Prosodic Focus in Hong Kong Cantonese without Post-focus Compression," in *Speech Prosody 2010*, Chicago. USA., 2010.
- [17] Y. Xu, "Post-focus compression: cross-language distribution and historical origin". *ICPhS. Hong Kong*, 152-155.
- [18] Qu, A. T. "The Language and Character of Tibetan". China Tibetology Press, 1996. (in Chinese)
- [19] F. Liu, "Intonation systems of Mandarin and English: A functional approach," Doctoral, Department of Linguistics, University of Chicago, 2009.
- [20] F. Liu, "Single vs. double focus in English statements and yes/no questions," presented at the Speech Prosody 2010, Chicago, USA, 2010
- [21] Z. J. Wang, Grammar of spoken Lhasa Tibetan (in Chinese): Minzu University Press, 1994.
- [22] Y. Xu, "ProsodyPro" available from, <http://www.phon.ucl.ac.uk/home/yi/tools.html>, 2005-2012.