Disentangling intonation in tone languages

Katrina Kechun Li, Francis Nolan, Brechtje Post

University of Cambridge
kl502@cam.ac.uk, fjn1@cam.ac.uk, bmbp2@cam.ac.uk

ABSTRACT

Describing intonation in tone languages can be difficult, as tone and intonation both heavily rely on f0. In this paper, we propose a method to disentangle intonation from tonal contours by plotting the f0 contour differences of the same sentence uttered under different focus conditions. We demonstrate this method through a case study of two Chinese languages, Chengdu and Changsha. The results reveal that while intonation realisation is not entirely independent of lexical tones, it is possible to identify a uniform picture of intonational realisation across various tonal conditions. This method lessens the reliance on prior knowledge of lexical tones in interpreting intonation, and it contributes to the development of an abstract representational system for intonation in tone languages.

Keywords: difference curve, tone languages, f0 contour, prosodic focus, Chinese languages

1. INTRODUCTION

Intonation and lexical tones are intertwined in tone languages, making it challenging to disentangle them. Consequently, studies of intonation in tone languages often require f0 contours to be interpreted separately for each tonal condition, which is less straightforward than in non-tonal languages. This limitation has hindered the application of an abstract intonational model in tone languages, even if ToBI-style models have been proposed [1]. This study examines a single intonational function, focus, in multiple tonal conditions, but explores a method that subtracts lexical tones from the intonation contour. Despite its simplicity, the visualisation yields a uniform picture of intonation patterning, which could support an abstract intonation model in tone languages.

2. METHODOLOGY

We present data from two Chinese languages with distinct tonal systems, Chengdu and Changsha. In each dataset, participants uttered a sentence with a comparable SVO structure as a response to a precursor wh-question eliciting focus on either the subject (Name) or the object (Object). Within the sentence, the Name, Verb, and Object all had the

same tone. Four sentences were constructed with four different tones (T1-T4) used at the three target positions. The example below shows one of the four tonal conditions in each language (citation tones represented in Chao number).

Name Prefix	Name	Verb	Aspect	Classifier	Object	Object Suffix
Chengdu						
ɕiaʊ53	**jɪn53**	**mai53**	lə0	pa53	**ji53**	tsɹ0
Xiao-yin		bought		a	chair	
Changsha						
ɕiaʊ42	**jɪŋ45**	**mɑi45**	tɑ21	tsɑ34	**iou45**	tsɹ33
Xiao-yin		sold		a	grapefruit	

The audio files were processed using *Praat* [2] and *Parselmouth* in Python [3], which generated a comparable data format that can be plotted using *ggplot2* in R [4]. For each syllable, f0 was extracted from 20 equidistant points of the rhyme and converted to semitones based on the participant's average f0.

For each dataset, we report the visualisations of four tonal conditions based on three repetitions of one male participant. The pattern revealed will become more obvious with a full dataset, but the current paper aims to demonstrate that even a small dataset can convey useful information through this visualisation. Information about the full dataset can be found in the congress paper 'Variations of focus prominence in three tone languages', and visualisations and sound samples are available on the website [5].

3. RESULTS AND DISCUSSIONS

The left panels of Figure 1 display the raw f0 contours for each syllable in Chengdu and Changsha respectively. Each row corresponds to a tonal condition, and each column corresponds to a syllable. It is notable that the f0 deviations happen mostly on the 'Name' and the 'Object' where the focus is placed (bold lines). However, the specific changes in f0 are heavily influenced by the underlying lexical tones, making it difficult to discern a clear pattern.

The difference curve introduced in the right panels of Figure 1, which equals the subject focus contour minus the object focus contour, presents a more consistent pattern across different tonal conditions. Instead of comparing two curves, one can compare a single curve with the reference line y=0 (red lines). Positive values indicate that the f0 of subject focus is higher than that of object focus, and vice versa.

The difference curves across tonal conditions seem to form groups. At the 'Name' in Chengdu, for example, T1 and T4 show a dipped difference curve below the reference line, whereas T2 and T3 have a curve above the reference line. This difference reflects the influence of underlying tonal target: the low target in T1 and T4 in the middle of the syllable gets lower under focus, and the high target in T2 and T3 gets higher, at the beginning and end of the syllable respectively. In Changsha, the difference curves cross the reference line in T1 and T2 at the 'Name', suggesting a bitonal target underlyingly, where the high target is scaled higher while the low target is scaled lower. This observation is in line with the proposal that focus is realised through hyper- or hypo-articulation of tonal targets [6].

In addition to the paradigmatic view of comparing to a reference line, the difference curve also promotes a syntagmatic view of comparison within the sentence. For example, the difference curve at the 'Object' can be close or above 0 (Chengdu T2/T4, Changsha T2/T3), which is not the expected focus effect. However, the relative difference between 'Name' and 'Object' still reflects the effect of focus. Although this relational view of intonation is not uncommon in non-tonal languages, it is often missing in the previous discussions in tone languages, due to the complexities in traditional visualisations.

In conclusion, the difference curve proposed in this study offers a simple yet effective way to disentangle intonation in tone languages. By visualising the f0 difference between subject and object focus, the method reveals a uniform picture on the effect of focus while taking into account the underlying lexical tones. The approach also renders easier both the paradigmatic and syntagmatic interpretation. The study shows that the proposed method can be used to generalise intonation patterns in different tonal languages. Future studies can explore the robustness of this method by examining its application on different sentence structures and tonal combinations.

5. REFERENCES

[1] Peng S., Chan M. K. M., Tseng C., Huang T., Lee O. J. and Beckman M. E. 2005. Towards a Pan-Mandarin System for Prosodic Transcription. In *Prosodic Typology*, edited by Jun S-A. Oxford University Press.

[2] Boersma P. and Weenink D. 2019. Praat: doing phonetics by computer. Accessed: May 26, 2019.

[3] Jadoul Y., Thompson B. and de Boer B. 2018. Introducing Parselmouth: A Python Interface to Praat. *Journal of Phonetics* 71, 1–15.

[4] Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

[5] Li K. K. 2023. InTone [Website]. Available: https://kc-li.github.io/InTone_Visualisation_Site

[6] Chen, Y. and Gussenhoven C. 2010. Post-Focus F0 Compression—Now You See It, Now You Don't. *Journal of Phonetics* 38 (4), 517–25.
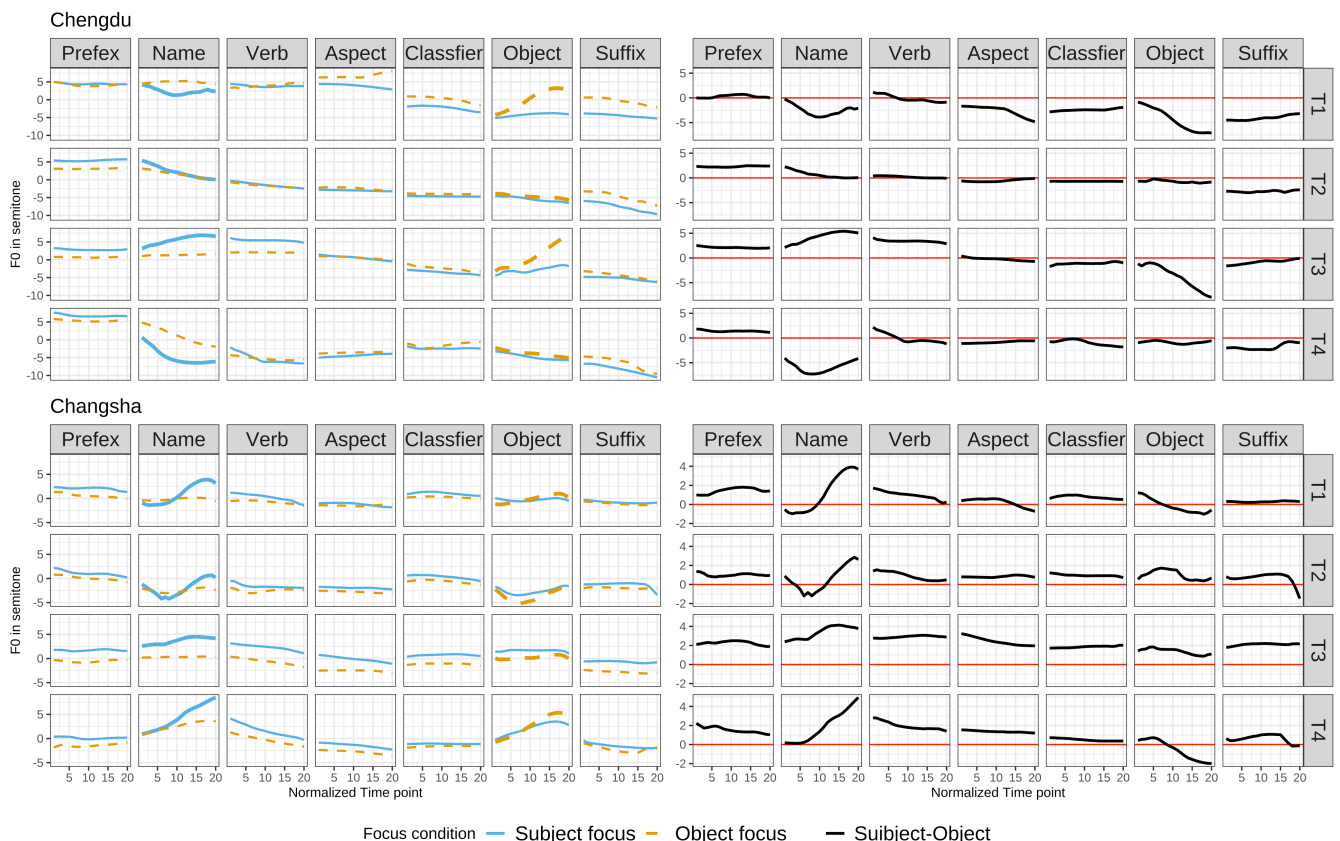
Figure 1: Raw f0 contour (left panel) vs. Difference curve (right panel), in Chengdu (top) and Changsha (bottom).