Untangling the Word-Tone System: 
The Basic Tonal and Prosodic Patterns in Choca-ngaca

Xiyao Wang

Department of Linguistics, University of Sydney, Australia
xwan5208@uni.sydney.edu.au

Abstract

In many Southeast Asian tone languages, the syllable comes as the domain of tone. However, a word-tone system has been found in several branches of the Bodic languages of the Himalayas and surrounding area. The Tamangic branch reported the most in literature, followed by Tibetan. Choca-ngaca is such a Tibetan example, yet it is severely underdocumented. This study makes the first attempt with one female speaker to investigate how word-tone works in the case of Choca-ngaca. Different lexical and phrasal contexts were checked for various syllabic constructions, i.e., monosyllable, disyllable, trisyllable and multisyllable. The results showed that tone in this word-tone language falls on the initial syllable of a word, and the high or low category of the tone leads the prosodic pattern of the entire word or phrase in the disyllabic context and beyond. The interaction between tone and consonant was also examined in a monosyllabic context, with a finding that Choca-ngaca tones contrast by the voicing condition of the obstruent initials and a pair of lateral approximant.

Index Terms: Choca-ngaca, word-tone, prosody, tone and consonant, Tibetan languages, Bodic languages

1. Introduction

1.1. The language: Choca-ngaca

Choca-ngaca is a Tibetan language, under the Bodic branch of Tibeto-Burman languages [3, 11]. It is spoken in the east of Bhutan by about 20,000 speakers, scattered among Mongar, Lhuentse, Trashi Yangtse and Trashigang districts [11]. ‘Choca-ngaca’ is an exonym, pronounced as /tʰoʔa-tɕa/ and alternatively spelled as ‘Chocha-ngachaka’ in [11]. The name means ‘you and us’, formed by chöca ‘you (pl.)’ and ngaca ‘us’, reflecting the native pronunciation for these words.

According to [11], Choca-ngaca has a pair of tonal contrasts, i.e., high vs. low registers. However, the tone system of the language is a type of word tone, which has been widely recognized in Tibetan [1, 9, 12], Tamangic [1, 2, 6, 7, 8, 10], Magaric [1] and East Bodish languages [4]. With relatively smaller speaking population, word-tone languages compared to the syllable-tone languages are often understudied, which can pose biases in the construction of tone typology. Choca-ngaca is such a case, with only one descriptive work [11], not to mention any work elaborating on its tone and prosody. The current study will thus make a descriptive contribution to this underdocumented language by demonstrating a basic profile of its tone and prosody, starting from monosyllables to multisyllables.

The sole basic description for Choca-ngaca does point out its tonal realization as word-tone, but no further account on how the system works in the language. This study will reveal some clues behind the puzzle and contribute to the understudied word-tone type. It also records an interaction between tone and consonant in Choca-ngaca’s tonal system, as the voiced obstruent onset is toned by a low register and the unvoiced by a high register. My study also investigates this feature, and the acoustic evidence is presented for the first time.

1.2. Word-tone

Word, rather than syllable, as the domain of tone in Lhasa Tibetan was noticed by Sprigg as early as in 1955 [8]. But the term ‘word-tone’ was not used until 1973 by Mazaudon for Tamangic languages [6]. By ‘word’, Mazaudon means ‘phonological word’, which can contain multiple syllables and morphemes, and where tone is not delimited by syllabic boundaries. The connotation has been continuously adopted [1], and Mazaudon noted again in 2014 that it is the whole word bears a single tone [8].

However, the word-tone does not run the same pattern across languages, such as the differences between Tibetan and Tamangic languages [7], and the different types within Tibetan languages [12]. For the former, the template case of Lhasa Tibetan has some extrametrical constituents, often unstressed inflectional suffixes, falling out of the phonological word domain, while in the Tamangic group, the suffixes do participate in the phonological word. For differences within the Tibetan group, there are three types of word-tone realization. Dolpo Tibetan does not see contrasts in tone register for non-initial syllables, while a level/falling contrast occurs for Lhasa Tibetan’s non-initial syllables, and the contrastive pitch registers and melodies can fall on the second syllables in Sherpa. But they all have four pitch contrasts on monosyllables which include contour tones. It is hence necessary to examine how the system operates in the case of Chocangaca that has only two level-tone contrasts.

2. Methods

2.1. Data recording

The linguistic data used in the study come from one young female speaker, who is native in Choca-ngaca and speaks the Tongshan dialect of Trashi Yangtse district. The context is selected as monomorphemic and dimorphemic words, and morphemic noun phrases. The elicitation materials were
designed according to the frame and list method introduced by [9] for Tamang’s tone. The carrier phrase for Choca-ngaca was finally adopted as (1) below, with reference to [4] for Kurtöp.

(1) nga khi ____ lap-mal
1.SG ERG ____ say-FUT
‘I will say ____’

The words and phrases were recorded in isolation for three times and in carrier phrase for once. But due to the lower quality of the recording of the tokens in frame, only the isolated tokens were used to display acoustic results.

2.2. Annotation and analysis

Before acoustic analysis, the recordings were annotated on Praat with three tiers – word, tone and gloss, as shown by the sample annotation in Figure 1:

![Sample Annotation on Praat](image)

Pitch values were then derived by using praatSauce (a Praat script developed by [5]), and imported into Excel. The plots for monosyllabic monomorphemic results were generated by R, and for the rest by Excel.

3. Results

3.1. The basic tone pattern of Choca-ngaca

3.1.1. Tonal contrasts and sonorants

By acoustic analysis, Choca-ngaca has two tonal contrasts, i.e., high- and low-level tones, which is clearly demonstrated by Figure 2.

![High and Low Tones in Choca-ngaca](image)

The F0 curves of the above graph were generated via the goem_smooth function in R, out of a set of minimal pairs with sonorant initials. The results showed that sonorant consonants in Choca-ngaca can be followed by both high and low tones. Table 1 below has presented the tokens used for analysis to exemplify the high and low tonal contrasts following the initial sonorants.

### Table 1: The high and low tones following initial sonorant consonants.

<table>
<thead>
<tr>
<th>Sonorant initial</th>
<th>High tone</th>
<th>Gloss</th>
<th>Low tone</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>/m/</td>
<td>mí</td>
<td>‘eye’</td>
<td>mì</td>
<td>‘person’</td>
</tr>
<tr>
<td>/n/</td>
<td>ná</td>
<td>‘nose’</td>
<td>ná</td>
<td>‘ná’</td>
</tr>
<tr>
<td>/ŋ/</td>
<td>ngá</td>
<td>‘drum’</td>
<td>ngá</td>
<td>‘I’</td>
</tr>
<tr>
<td>/ɲ/</td>
<td>ŋyá</td>
<td>‘to borrow’</td>
<td>ŋyá</td>
<td>‘fish’</td>
</tr>
<tr>
<td>/l/</td>
<td>lé</td>
<td>‘tongue’</td>
<td>lé</td>
<td>‘karma’</td>
</tr>
</tbody>
</table>

3.1.2. Tone and voicing

In contrast to sonorants, tone performs differently with the obstruent initials. The low tone only comes with the voiced obstruents, and the high tone with the voiceless, as exemplified in Table 2. The high and low distinctions also apply to the voiced vs. voiceless approximant initials, i.e., là vs. ħà, also displayed in Table 2.

### Table 2: The high and low tones following voiced vs. voiceless initials.

<table>
<thead>
<tr>
<th>Initial [+voice]</th>
<th>High tone</th>
<th>Gloss</th>
<th>Initial [+voice]</th>
<th>Low tone</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>/pʰ/</td>
<td>pʰú</td>
<td>‘to blow’</td>
<td>/b/</td>
<td>bú</td>
<td>‘insect’</td>
</tr>
<tr>
<td>/p/</td>
<td>pú</td>
<td>‘body hair’</td>
<td>/b/</td>
<td>bú</td>
<td>‘insect’</td>
</tr>
<tr>
<td>/tʰ/</td>
<td>tʰú</td>
<td>‘to pick up’</td>
<td>/d/</td>
<td>dū</td>
<td>‘to sit’</td>
</tr>
<tr>
<td>/t/</td>
<td>tú</td>
<td>‘to wash’</td>
<td>/d/</td>
<td>dū</td>
<td>‘to sit’</td>
</tr>
<tr>
<td>/kʰ/</td>
<td>kʰé</td>
<td>‘ice’</td>
<td>/k/</td>
<td>ké</td>
<td>‘where’</td>
</tr>
<tr>
<td>/k/</td>
<td>ké</td>
<td>‘ice’</td>
<td>/k/</td>
<td>ké</td>
<td>‘where’</td>
</tr>
<tr>
<td>/ɡ/</td>
<td>ɡè</td>
<td>‘to step over’</td>
<td>/ɡ/</td>
<td>ɡè</td>
<td>‘to step over’</td>
</tr>
<tr>
<td>/s/</td>
<td>sá</td>
<td>‘soil’</td>
<td>/z/</td>
<td>zá</td>
<td>‘to eat’</td>
</tr>
<tr>
<td>/c/</td>
<td>cá</td>
<td>‘flesh’</td>
<td>/z̥/</td>
<td>z̥à</td>
<td>‘chicken’</td>
</tr>
<tr>
<td>/t/</td>
<td>t̥ú</td>
<td>‘to vomit’</td>
<td>/d̥z̥/</td>
<td>d̥z̥à</td>
<td>‘to run’</td>
</tr>
<tr>
<td>/ts/</td>
<td>tsá</td>
<td>‘rook’</td>
<td>/t̥s̥á/</td>
<td>t̥s̥á</td>
<td>‘rook’</td>
</tr>
<tr>
<td>/tʃ/</td>
<td>tʃá</td>
<td>‘grandson’</td>
<td>/tʃá/</td>
<td>tʃá</td>
<td>‘grandson’</td>
</tr>
</tbody>
</table>

The F0 curves of the above tokens were drawn via R in Figure 3 below, which provides acoustic evidence to [11]’s account for the voicing condition. The voiced obstruent initials do show a lower F0 register than the unvoiced equivalents, so as the voiced lateral approximant.

![High and Low Tones with Voiced vs. Voiceless Initials in Choca-ngaca](image)
Among the unvoiced plosives, it seems that the aspirated initials gain a higher F0 register to follow, rather than the unaspirated one as recorded in [11].

3.2. The prosody beyond monosyllables

3.2.1. Disyllabic patterns

The F0 curves of the disyllabic monomorphemic words have been demonstrated in Figure 4.1a, b. In overall, the high-toned words have a higher pitch register than the low-toned words. And a distinctly higher pitch falls on the initial syllables of the high-toned words (e.g., tò in tòwa, ‘stomach’ vs. dò in dòwa, ‘stone’ in Figure 4.1a). This is consistent for the syllable structures of CVCV (i.e., tòwa, ‘stomach’ and kíli, ‘elbow’ vs. dòwa, ‘stone’ and kúto, ‘head’1) and CVCCV (i.e., kʰínto, ‘roof’ vs. jáŋko, ‘chest’). However, the pitch is neutralized on the second syllables of all high- and low-toned words.

Notes:

b. kʰínto, ‘roof’; jáŋko, ‘chest’

4.1 Disyllabic monomorphemes

Notes:

a. tɕʰó-i, ‘your’; kʰó-i, ‘his’; tsó-i, ‘lake’s’
gó-i, ‘my’; mò-i, ‘her’; là-i, ‘mountain pass’s’
b. tɕʰó-tɕa, ‘you (pl.)’; tsó-ɡe, ‘in lake’
gó-ɡe, ‘we’, là-ɡe, ‘in mountain pass’

1 This low-toned disyllabic monomorphemic word has an unvoiced obstruent initial, which runs against the results of the monosyllabic words. I hence propose that the voicing condition for tone only applies to the monosyllable, and a recent tonogenesis might have appeared on the disyllabic monomorphemic words, which needs further study in the future.

4.2 Disyllabic dimorphemes

The disyllabic dimorphemic words have the same pattern with the disyllabic morphemic words, as illustrated in Figure 4.2a, b. c. Across the three kinds of syllable structures (i.e., CV-V, CV-CV, CV-CVC), the first morphosyllables of the high-toned words have a higher register than that of the low-toned ones. And the second morphosyllables of all the disyllabic dimorphemic words tend to be neutralized. This can be exemplified by the words tɕʰó-tɕa (‘you pl.’), tsó-ɡe (‘in lake’) vs. gá-ɡa (‘we’), lá-ɡe (‘in mountain pass’) in Figure 4.2b that have a CV-CV structure. The roots tɕʰó (‘you’) and tsó (‘lake’) in the former shows a higher register than gá (‘I’) and lá (‘mountain pass’), but the suffixes -tɕa (plural) and -ɡe (locative) do not show distinctive pitch patterns between the high- and low-toned words. The results are the same for the remaining word samples with the CV-V structure in Figure 4.2a and CV-CVC in Figure 4.2c.

Pulling mono- and di-morphemes together, the prosodic pattern of the disyllabic context in Choqa-ngaca tends to be determined by the tone type of the initial syllables. The words with a high-toned initial syllable have a higher register than those with a low-toned counterpart, regardless of syllable structure.

3.2.2. Trisyllabic pattern

The pitch traces of trisyllabic words with the structures of CV-CV and CV-CV-CV are shown in Figure 5a, 5b. The words in a trisyllabic context tend to show the same pattern as in the disyllabic context. When the initial morphosyllable of the trisyllabic word has a high tone register by its own, the prosodic pattern of the complete word has a relatively higher register, with a neutralization effect on the remaining two morphosyllables.

Notes:

a. tɕʰó-tɕa-i, ‘yours (pl.)’; gá-ɡa-i, ‘ours’
b. kʰó-raŋ-ja, ‘himself also’; kʰó-raŋ-ɡi, ‘it is him only’
mò-raŋ-ja, ‘herself also’; mò-raŋ-ɡi, ‘it is her only’

Notes:

a. tɕʰó-tɕa-i, ‘yours (pl.)’; gá-ɡa-i, ‘ours’
b. kʰó-raŋ-ja, ‘himself also’; kʰó-raŋ-ɡi, ‘it is him only’
mò-raŋ-ja, ‘herself also’; mò-raŋ-ɡi, ‘it is her only’
Figure 5: Trisyllabic tone patterns in Choca-ngaca.

Take a pair of words from Figure 5b for example. The root kʰó ('he') has a high tone as a monosyllabic monomorpheme alone, and mò ('she') with a low tone. A higher register remains with the former root in the trisyllabic trimorphemic construction of the word kʰó-ray-gi ('it is him only'), whereas a lower register with the latter in mò-ray-gi ('it is her only'). Except for the initial morphosyllable, the remaining suffixes -ray (reflexive) and -gi (genitive) all carry a non-contrastive pitch register, regardless of the high- or low-toned initial roots they are following. The results show that it is still the initial syllable that determines the whole prosodic pattern for the trisyllabic context, no matter how many syllables constitute this word.

3.2.3. Multisyllabic pattern

Noun phrases were selected to check Choca-ngaca’s prosodic pattern beyond the word domain. They are comprised of two words, a possessor plus the head noun, with a disyllabic monomorphemic and dimorphemic words. Similar registers were attained on the second syllables of all disyllabic monomorphemic and dimorphemic words.

<table>
<thead>
<tr>
<th>Tone Pattern</th>
<th>Example Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>'H+H'</td>
<td>khó-i kùto</td>
</tr>
<tr>
<td>'H+L'</td>
<td>raŋ-úgu</td>
</tr>
<tr>
<td>'L+H'</td>
<td>tɕʰ-úgu</td>
</tr>
</tbody>
</table>

Figure 6: Multisyllabic tone patterns in Choca-ngaca.

The tone types of the second word in each phrase do not make that of the initial word changed, although they do exert certain influence on the registers of the initial within their categories. The high tone remains high before both high and low tones, and the low tone remains low under the same condition. But the high-toned initial word followed by another high-toned word has a higher register than those with a low-toned word (e.g., tɕʰ-o-i úgu vs. tɕʰ-o-i kùto). And the low-toned initial word coming with another low-toned word has the lowest register (e.g., ɲà-i úgu vs. ɲà-i kùto).

Similarly, the second word is also affected by the initial word in each phrase. The high tone remains high after both high and low tones (e.g., tɕʰ-o-i úgu vs. ɲà-i úgu), and the same for the low tone (e.g., tɕʰ-o-i kùto vs. ɲà-i kùto). However, when the high-toned word is initialled by a low-toned word, it has a lower register than that initialled by another high-toned word. For the low-toned word, there is no consistent pattern to show how the low tone is affected by the category of the initial tones, out of the two pairs of samples, i.e., tɕʰ-i kùto, kʰó-i kùto vs. ɲà-i kùto, mò-i kùto.

The difference is that the influence of the first word on the second falls on the initial syllable of the latter, but the effect the other way round falls on both morphosyllables (root and suffix) of the first word. Its initial morphosyllable (root) still determines the tone types, but the register varies within the category. And its second morphosyllable (suffix) does not show a neutralized pitch as the prosody of the phrase continues on the following word. In contrast, neutralization occurs on the second syllable of the second word, which is also the final syllable of the whole phrase.

In sum, the prosodic pattern of multisyllabic context is primarily determined by the tone type of the initial word, or more precisely, its initial syllable, while also affected by the tone type of the second word and the initial syllable of it. The tone category of the word in a phrasal context stays notably stable, whether initialled or followed by another toned word. This proves that the domain of tone in Choca-ngaca does not extend to the phrase but retains in word. The neutralization effect is only realized on the last syllable of the whole phrase, which is similar to the case when the effect applies to word.

4. Discussions

The research results showed that the high and low-level tones, as the mere two tonal contrasts of Choca-ngaca, perform differently with different types of consonant. They can both fall on the vowels following sonorant consonants. However, there is a voicing condition applied to the tone following obstruct consonants and a pair of lateral approximant. The voiced initials are generally identified with low tone, and the voiceless with high tone. A contrast of tones in terms of aspiration also occurred among the unvoiced consonants. The aspirated initials have a higher register than the unaspirated one within the high tone category, which is not consistent with the findings in [11].

For disyllabic words, the whole prosodic pattern is determined by the initial syllable. The words with a high-toned initial syllable have a higher register overall, regardless of syllable structure. And the initial syllables of the high-toned words all carry a distinctively higher pitch than that of the low-toned words. However, the pitch contrast between the second syllables of high- and low-toned words is neutralized — very similar registers were attained on the second syllables of all disyllabic monomorphemic and dimorphemic words.

Trisyllabic words perform very similar to the disyllabic in prosodic patterns. The words with high tones have a higher register than those with low tones in overall. Higher pitch is overwhelmingly applied to the initial syllables of the former than that of the latter. But again, a neutralization effect is imposed on the non-initial syllables across all trisyllabic words. Gathering the results of di- and tri-syllabic words together,
Choca-ngaca’s word prosody in the context of over one syllable is very likely to depend on the pattern of the initial syllable, regardless of the number of the syllables that constitute this word.

In the phrasal context, whether the word constituents are high- or low-toned, they retain the same tone type as they are alone. But they can be affected by each other within the tone category. A higher register occurred on the initial words of the phrases that are high toned than that are low toned. But the register of the low-toned word can be lowered when followed by another low-toned word, and that of the high-toned word can be raised when followed by a second high-toned word. And the neutralization effect does not apply to the secondary syllable of the initial word, when it is in the medial position of the phrase. But the effect falls on the second syllable of the second constituent, i.e., the last syllable of the phrase. To sum up, the overall prosodic pattern of multisyllabic phrases is primarily determined by the pitch of the initial syllable of the initial word constituent, while also affected by the second constituent. The constituents retain their own tone types, which means that the domain of tone in Choca-ngaca stays within word. They manipulate each other’s register within the tone category, so as to cooperatively attain a phrasal prosodic profile.

5. Conclusions

This study has investigated the tone pattern of the underdocumented Choca-ngaca with acoustic evidence. The language has a word-tone system, with a contrast of high and low registers. The initial syllable is the bearing unit of the word tone, but the tone type applies to the whole word. The pitch of the initial syllable determines the entire prosodic pattern of a word or phrase, whether in a disyllabic context or beyond. But the word constituents in the phrase context retain their tone categories as alone, so the domain of tone does not stretch beyond the word. In addition, consonantal laryngeal categories interact with tone in monosyllabic context, regarding the voicing and aspiration conditions. Given that only the isolated environment was checked, the findings in the present study needs to be further examined under unisolated condition. Future studies should also include a greater number of speakers, across different ages and genders. And statistical study also needs to be conducted in the future with more words and phrases, in order to confirm the prosodic patterns revealed in this study.

6. Acknowledgements

I would like to extend thanks to my PhD supervisor Dr. Gwendolyn Hyslop, who invited the participant for this study and helped me along the way. I also want to thank Mr. Carl Bodnaruk, who helped me with the minimal pairs and data recording.

7. References


