Preservation of lexical tones in singing in a tone language

Anastasia Karlsson1, Håkan Lundström2, Jan-Olof Svantesson1

1 Dept. of Linguistics and Phonetics, Centre for Languages and Literature, Lund University, Sweden
2 Malmö Academy of Music, Lund University, Sweden
anastasia.karlsson@ling.lu.se, hakan.lundstrom@kanslik.lu.se, jan-olof.svantesson@ling.lu.se

Abstract

Lexical tones are important for expressing meaning and usually have high priority in tone languages. This can create conflicts with sentence intonation in spoken language and with melodic templates in singing since all of these are transmitted by pitch. The main question in this investigation is whether a language (in our case the Mon-Khmer language Kammu) with a simple two-tone system uses similar strategies for preserving lexical tones in singing and speech. We investigate the realization of lexical tones in a singing genre which can be described as recitation based on a partly predefined, though still flexible, melodic template. The contrast between High and Low tone is preserved, and is realized mainly at the beginning of the vowel. Apparently, the rest of the syllable rhyme serves other purposes. Syllables are often reduplicated in singing, and the reduplicant ignores lexical tones. The preservation of lexical tones in Kammu singing, and their early timing close to the vowel onset, is very similar to what we have found for speech.

Index terms: singing, tone languages, Mon-Khmer

1. Introduction

1.1. Lexical tones and music

The study of the interaction between music and language is a small but rather active branch of musicology. Chinese singing and recitation has drawn the attention of several scholars, among others Chao [1], who defines different singing styles with differing relationships between pitch and lexical tone. Thai song has been studied by among others Morton [2] and Mendenhall [3]. In Chinese, Thai and other tone languages in Southeast Asia which have several contour tones, the relation between pitch and lexical tone is normally rather complex and not always predictable by simple rules; see for example Tanase-Ito [4] concerning Thai court song. Concerned with Mandarin Chinese, Wee [5] sets up rules that take into account salient parts of the lexical tones and their combination with musically stressed tones. Oesch [6] concludes that in Yoo tradition in Thailand, “if the song is syllabic … the level of the musical version is defined by the relative position of a tone to its preceding tone” and “if the song is melismatic … the intonation of the word is expressed in music by an ascending or descending melismatic configuration”. He also notes that final formulae are musical culminations where the musical movement often dominates. Schellenberg [7] examined the realization and perception of lexical tones in Cantonese and Mandarin singing. Using especially composed songs he found that Mandarin singers and listeners do not incorporate lexical tones missing from written music, while Cantonese singers add extra tonal information not found in the written music, using this information for understanding the sung words.

1.2. Kammu speech

In this article we examine if and how lexical tones are preserved in singing using data from Kammu, a Mon-Khmer language spoken mainly in Northern Laos. Kammu is a tone language of the Southeast Asian type, where each syllable has a tone, either High or Low. Compared to Cantonese, Thai and Mandarin, Kammu has a less complex tone system, with two lexical tones and no contour tones. We know from investigations of spoken Kammu that it is important to preserve the identities of the tones and we believe that this should apply for Kammu singing as well.

The Kammu tones have rather high functional load (more than 900 minimal pairs have been found for the tone contrast) and are clearly distinguished by tonal speakers despite the rather small distance between them [8]. The tones are rather level and well preserved in speech. The relation between adjacent tones is also important to signal tonal identities. We found that the lexical tones overrule phrase intonation (high phrase boundary at the right phrase edge) if the final words have the tone combination High + Low tone. In this case, the high phrase boundary tone is not fully realized since it might jeopardize the identity of the final Low tone. This shows that lexical tone has priority over intonation in Kammu [9]. We also found that tones tend to be realized as close the onset consonant as possible [9].

Based on these findings we can make some predictions about the realization of tones in singing. We predict that if tones are preserved, F0 will be the strongest cue, and also that tones will be realized early in the rhyme, close to the onset consonant. Lexical tones can either be overruled by the melodic template or they can be preserved, in which case the melody will be realized in the later part of the syllable, similar to what we found for the realization of sentence intonation in speech. In the Kammu singing genre chosen for the present investigation, the singer may elaborate the words by prolonging or contracting them. As Kammu has a contrast between long and short vowels, the words cannot be prolonged freely, so therefore we will also investigate how the prolongations are achieved. There is a lot of lengthening and reduplication of syllables in Kammu singing and we predict that the reduplicants ignore lexical tones and are mainly used for the realization of melody.
1.3. Kammu tə́ ə̀m singing

In earlier publications [10, 11], we have shown that the vocal tradition of the Kammu can be seen as a mono-melodic system in which a large number of orally transmitted poems, trnə́om, are sung according to a limited number of melody types, varying with the situation. The singing is orally transmitted and each performance constitutes a recreation of the trnə́om which includes a certain amount of variation both with regard to the music and the poetry as well as the way the poetry and the basic melodic structure are adapted to each other.

Of these singing styles tə́om is the most complex. Tə́om singing is basically solo singing, but it plays an important role in social situations where a number of individuals will take turns singing alternating songs on the same basic melody. The singers may embellish the poems with words of praise and politeness and also add sets of words that may be traditional or made up on the spot.

The melodic template of tə́om varies between the Kammu dialect areas, but also in details between villages and even between individuals. For this study two melodies of neighbouring Kammu dialect areas in northern Laos have been used, namely those of Kwɛ̀ɛ̀ and Yūan. They share the same basic outline but differ in details of the initial and final formulae. While the remaining part – the recitation section – of the former is based on an even rhythm and two musical pitches, the latter utilizes an iambic pattern and three pitches.

In tə́om singing, the relation between pitch and lexical tone is not clear-cut and simple. In parts of the singing, particularly in initial and final formulae of phrases, musical factors dominate over the lexical tones (music-pitch centration) and in others the lexical tones dominate (lexical tone centration). It can also be anticipated that differences in our results to some degree may reflect the differences between the two melodic templates used.

Prolongation of syllables is a characteristic feature of tə́om singing. It takes two different forms, either as reduplication of the vowel (this type of prolongation does not occur in spoken Kammu) or as lengthening of a final sonorant, e.g.:

\[
\begin{align*}
\text{CVC} & > \text{CV.CV}: \quad \text{ʔə́m} > \text{ʔə́.tə́m} \\
\text{CV(V)} & > \text{CV.hV(C)}: \quad \text{tāa} > \text{ta.ha} \\
\text{CVVC} & > \text{CV.CV}: \quad \text{jə́a} > \text{ja.ʔa} \\
\text{CVS} & > \text{CV.SV} \quad \text{ʔə́om} > \text{ʔə́.mə} \\
\text{CV(V)S} & > \text{CV(V).S}}: \quad \text{nə́an} > \text{nə́.ŋə́n}
\end{align*}
\]

Here, S is a sonorant and final C is a non-sonorant consonant (stop or fricative); VV is a long vowel or a diphthong.

2. Material

The material used for this investigation is part of a larger corpus of Kammu singing and music from northern Laos and Thailand recorded in the 1970s by Kristina Lindell and later by one of the authors (HL) with the purpose of investigating Kammu song and music in general.

Singer 1: Sák (male), Rmčial village, Yūan dialect area, Laos; recorded in a session including four male singers in a private home probably in northern Thailand in 1974 using the Kwɛ̀ɛ̀ area melody. The title of the song is Táa píc ʔāaj jiə́ priə́aj “Do not abandon me to stay with others”, about one minute recording.

Singer 2: Kàm Rǎw (male), Rmčial village, Yūan dialect area, Laos; recorded in a session including eight male singers during field work in northern Thailand, autumn 1981, using the Yūan area melody. The title of the song is Táa píc ʔāaj jiə́ priə́aj “Do not abandon me to stay with others”, about one minute recording.

Singer 3: Nàaŋ (female), Mọ̀ Lɔ̃xt village, Yūan dialect area, Laos. Recorded at a farewell party in Lampang, Thailand, 19 July 1974, using the Kwɛ̀ɛ̀ area melody. The title of the song is Kɔ́y miə́j kʰɔ́y liə́ŋ kə́mi? “Be careful to remember your dear one back here”, about 2.5 minutes recording.

3. Analysis

For each syllable, F0 was measured at the vowel onset and at the end of the next turning point of the pitch movement. These F0 values were measured both in the original word and in the reduplicant (the reduplicated vowel or the prolonged sonorant). All measurements were done using Praat [12]. Two points were always taken irrespective of the number of turning points within the syllable. Many syllables have more than two turning points of pitch, this is especially characteristic for considerably prolonged and especially reduplicated syllables. These F0 ripples result from the singer’s vibrato and belong to the musical performance. They are found for two of the three singers. In cases with more than two turning points, the second F0 value was measured in the first turning point after the vowel onset. All measurements were performed manually.

An example of vibrato and measurement points is given in Figure 1. The word is sii-pāaj ‘bean’. The first syllable sii has High tone, and the second syllable has Low tone and is reduplicated: pāaj > pāaj.a. Vibrato occurs in the reduplicant ja. f01 and f02 indicate the points of measurement in each syllable: f01 is measured at the vowel onset, and f02 is measured at the next F0 turning point. The initial pitch fall in sii is typically found for words with High tone, and the rise in pāaj for words with Low tone.

![Figure 1. Illustration of vibrato and measurement points of F0. The word is sii-pāaj ‘bean’ with reduplicated pāaj. Sii.pāaj.aa. Singer 2.](image)

4. Results

For Singer 1 (Sák), 78% of all syllables are prolonged (reduplicated vowel or lengthened sonorant coda). Singer 2 (Kàm) has 39% prolongation and Singer 3 (Nàaŋ) 68%.

The results of the F0 measurements are shown in Table 1 and Table 2. In Table 1, F0 values in syllables and reduplicants are shown (in semitones relative to 100 Hz).
Table 1. F0 values in syllables and reduplicants (semitones relative to 100 Hz)

<table>
<thead>
<tr>
<th></th>
<th>syllable</th>
<th></th>
<th></th>
<th></th>
<th>reduplicant</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>begin</td>
<td>end</td>
<td>n</td>
<td>begin</td>
<td>end</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>Singer 1</td>
<td>mean F0 (st)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High tone</td>
<td>9.08</td>
<td>10.07</td>
<td>11</td>
<td>7.64</td>
<td>7.02</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low tone</td>
<td>5.67</td>
<td>6.73</td>
<td>56</td>
<td>6.44</td>
<td>6.05</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>difference</td>
<td>3.41</td>
<td>3.34</td>
<td></td>
<td>1.20</td>
<td>0.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-test (p-value)</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singer 2</td>
<td>mean F0 (st)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High tone</td>
<td>10.27</td>
<td>6.61</td>
<td>23</td>
<td>7.90</td>
<td>6.88</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low tone</td>
<td>3.48</td>
<td>4.74</td>
<td>38</td>
<td>4.53</td>
<td>4.42</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>difference</td>
<td>6.79</td>
<td>1.87</td>
<td></td>
<td>3.37</td>
<td>2.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-test (p-value)</td>
<td>&lt; 0.001</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>0.006</td>
<td>n.s.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singer 3</td>
<td>mean F0 (st)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High tone</td>
<td>18.21</td>
<td>16.52</td>
<td>28</td>
<td>16.15</td>
<td>15.60</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low tone</td>
<td>14.65</td>
<td>16.35</td>
<td>69</td>
<td>15.05</td>
<td>15.17</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>difference</td>
<td>3.56</td>
<td>0.17</td>
<td></td>
<td>1.10</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-test (p-value)</td>
<td>&lt; 0.001</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. F0 fall in syllables

<table>
<thead>
<tr>
<th></th>
<th>Singer 1</th>
<th>Singer 2</th>
<th>Singer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High tone</td>
<td>Low tone</td>
<td>High tone</td>
</tr>
<tr>
<td></td>
<td>mean fall (st)</td>
<td>-0.99</td>
<td>-1.06</td>
</tr>
<tr>
<td></td>
<td>t-test (p-value)</td>
<td>n.s.</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

For all singers there is a highly significant (p < 0.001) F0 difference between High and Low tone at the beginning of syllables. The difference is about 3–4 semitones, comparable to the difference found in Kammu speech [8]. There is a difference towards the end of the syllable as well, but for Singers 2 and 3 it is much smaller and not significant. In the reduplicant, syllables with High tone have higher F0 on the average, but significance tests show that this difference is not upheld consistently (except for the beginning of the reduplicant for Singer 2).

We also tested if there is a fall or a rise in each syllable with a paired t-test. The general pattern is that there is a fall in syllables with High tone and a rise in syllables with Low tone. The only exception is Singer 1 whose High tones have a rise, which is, however, not significant. This supports our hypothesis that the beginning of the syllable carries the tone in tiöm singing: F0 falls from a high to an intermediate neutral position if the tone is High, and rises from a low to a neutral position if the tone is Low.

5. Discussion

High and Low lexical tones are kept apart in both spoken and sung Kammu. As is the case in speech, tones are more likely to be realized near the vowel onset. Tones in Kammu developed from the loss of contrast between initial voiced and voiceless consonants and can still be seen as a feature of the initial consonant [8]. This may explain the prevalence for the early timing of the tonal contrast in both speech and singing.

The following part of a syllable does, however, not completely ignore tones. Singer 1 still shows a significant difference for the tone contrast in the later part of the vowel. Singers 2 and 3 show an interesting relation between the type of tone and pitch movement within the syllable. Thus, at High tone pitch tends to fall within the syllable, and at Low tone pitch is often rising. This can be explained either as a way to increase the auditory difference between the tones or by the singers’ striving towards a neutral pitch level. This is also similar to what was found for speech. The tonal peak in CV(V)N (N = nasal) words with Low tone has a later timing than the tonal peak in CV(V)N words with High tone, which often results in a slightly rising movement within words with Low tone [9].

Similar to what takes place in speech, sonorant codas are prolonged in singing, the difference being that in speech we do not find any considerable pitch movements within sonorant codas, while in singing the movement is larger, and sonorant codas can function as syllables of their own.

In singing, syllables are often reduplicated. This is done to fit the rhythmic template but probably also for the realization of melodic templates without jeopardizing the identities of the lexical tones. Reduplicants ignore lexical tones. It is interesting that Singer 2 shows some significance between tone type and the reduplicant.

6. Conclusion

Tones are kept apart in Kammu singing. Kammu has a two-tone system with level tones and the general pattern for the singers is to uphold the contrast in the beginning of syllables, while using the rest for other purposes. Reduplication of syllables is a way to realize the melodic template while avoiding conflicts with the lexical tones. The pattern found for Kammu differs from what has been found for singing in Southeast Asian languages with complex tone systems and we believe that contour tones have rather different principles of interaction with intonation and melodic pitch than the level tones in languages like Kammu.

7. Acknowledgments

The work reported here was done within the project In the borderland between song and speech, financed by the Swedish Research Council.
8. References


