Evidence for contrastive tonal alignment in Shilluk

Bert Remijsen, Otto Gwado Ayoker

University of Edinburgh, United Kingdom
Corresponding author email: b.remijsen@ed.ac.uk

Abstract

It is widely assumed that tonal alignment is not contrastive in contour tones [1,2,3]. Contrary to this consensus, [4] recently presented evidence of contrastive alignment in falling contours in Dinka. The current study contributes to this area of investigation through an acoustic analysis of the relevant part of the tone system of another language: Shilluk. The dataset is built around the tonal contrast of Low vs. Early-aligned High Fall vs. Late-aligned High Fall vs. High. The descriptive statistics resulting from this study are in line with the hypothesis that tonal alignment can indeed be contrastive in falling contour tones.

Index Terms: tone, tonal alignment, contour tones, Shilluk

1. Introduction

1.1. Motivation

Tonal alignment refers to the timing of changes in fundamental frequency (f0) relative to the segmental sequence. This timing relation determines how a tone pattern is perceived. The same pattern of f0 change can give rise to different pitch impressions, depending on its alignment. This is illustrated schematically in Figure 1. The four traces have the same shape: high f0 is followed by a drop to a lower level. The pitch impression on the central CV syllable depends on the alignment of this f0 change. When the f0 drop is largely completed before the vowel of the central syllable, as in the leftmost pattern, then the central syllable is prone to be perceived with low pitch. In contrast, if the f0 drop is near the end of the central syllable, as in the rightmost pattern, then the central syllable will be perceived with high pitch. When the drop is aligned early in or at the center of the vowel, as in the two patterns in the middle, then the result is a falling f0 contour over the vowel, the part of the syllable with greatest intensity, so that the falling contour is likely to perceptually salient in some way. In the phonological representations of such patterns, two targets – high and low – are associated in some way with the same syllable. Such patterns are known as contour tones.

The question at issue here is whether a language can have two falling contour tones, as in Figure 1. It is widely hypothesized that tonal alignment is not contrastive in contour tones [1,2,3]. Odden writes: “[I]t might be that in some languages pitch changes are timed relatively early in the syllable, and in other languages they are timed relatively late. Such control would only be phonetic, never phonological” [2:450]. Hyman postulates a rule to ensure that “there is no possible opposition between two HL or two LH contours where the two tones are synchronized differently within the syllable.” [1:51]

However, a recent study on Dinka, a Nilo-Saharan language spoken in South Sudan, presents evidence for precisely such a configuration, of a contrast between early- vs. late-aligned falling contours [4]. F0 traces illustrating the contrast between the two contours are presented in Figure 2, in the context of a non-high target within the same word. Separate traces are displayed for realisations on syllables with a short (CVC, full line) and a long vowel (CVVC, interrupted line), respectively. These results show that tonal alignment is distinctive in a consistent manner, both when time pressure is smaller (CVVC) and when it is greater (CVC).

However, while Dinka has both Low and Early-aligned Fall tone patterns, these two configurations are in complementary distribution; they do not represent distinct lexical or morphological specifications for tone.

If tonal alignment is indeed potentially contrastive, then it should be possible for a language to contrast all four of the patterns shown schematically in Figure 1. Here the two falling contour tones stand in contrast not only with one another but also with level tone patterns that are similar in terms of tonal alignment. The goal of this study is to present new evidence on this question, from the tone system of Shilluk. As in earlier work on tonal alignment [4,5,6,7], vowel length is controlled so as to examine alignment across levels of time pressure.

Figure 1. Schematic representations of four falling fundamental frequency patterns over a 3-syllable domain.

Figure 2. Averaged f0 traces on a normalised time axis, for Early-aligned (blue) vs. Late-aligned (red) falling contours in Dinka. Separate traces by vowel length: short (full) vs. long (interrupted). The traces show descriptive statistics across 4 segmental sets and 13 speakers. From [4].
1.2 Background on Shilluk

Like Dinka, Shilluk is part of the Western Nilotic group within the wider Nilo-Saharan language family. A crucial characteristic in relation to the current study is that content-word roots are primarily closed monosyllables [8]. These roots can combine with affixes, but there is also a rich system of stem-internal marking [9]. Tone and vowel length in particular are extensively used in morphological paradigms.

The tone system includes a total of seven contrastive tone categories, all of which occur productively as transitive verb inflections, as seen from Table 1. There are the level patterns Low (cv´c), Mid (cv´c), and High (cvc); the Rise (cv´c); and three falling categories. Two of these have as the initial target of the contour the same level as the High tone: the Early High Fall (cv´c) and the Late High Fall (cv´c). There is also a Low Fall (cvc), which falls from the level of the Mid tone. [NB The transcription of Late fall (cv´c) is distributed over vowel and coda. This is as an ad hoc answer to the need to distinguish all categories, and does not prejudge the phonological analysis.]

Table 1. A minimal-set illustration of the seven Shilluk tone categories. Each category is illustrated by an inflection selected from the paradigms of two transitive verbs. The target syllable is in bold face.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Examples (sounds are embedded, click to play)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>jįǹ jˇiŋ ki kˇjŋ</td>
</tr>
<tr>
<td></td>
<td><em>You have taken money here.</em></td>
</tr>
<tr>
<td>Mid</td>
<td>friend:j Past-take:APPL money: happiest</td>
</tr>
<tr>
<td></td>
<td>‘Somebody used a friend to take money.’</td>
</tr>
<tr>
<td>High</td>
<td>stick:j Past-beat:APPL:2SG drum: happiest</td>
</tr>
<tr>
<td></td>
<td>‘You have used a stick to beat the drum.’</td>
</tr>
<tr>
<td>Low Fall</td>
<td>drum:j Past-beat:APPL:2SG PRP place:S.DEF</td>
</tr>
<tr>
<td></td>
<td><em>You have beaten a drum here.</em></td>
</tr>
<tr>
<td>Early</td>
<td>drum:j Past-beat PRP place:S.DEF</td>
</tr>
<tr>
<td></td>
<td><em>Somebody has beaten a drum here.</em></td>
</tr>
<tr>
<td>Late High</td>
<td>jįǹ jˇiŋ pάc</td>
</tr>
<tr>
<td></td>
<td>money:j Past-take:FG village:2SG</td>
</tr>
<tr>
<td></td>
<td>‘Somebody went to the village to take money.’</td>
</tr>
<tr>
<td>Rise</td>
<td>friend:j Past-take:APPL:2SG money: happiest</td>
</tr>
<tr>
<td></td>
<td>‘You used a friend to take money.’</td>
</tr>
</tbody>
</table>

Aside from tone, vowel length is also contrastive in Shilluk. Just as in related Dinka [4], it is contrastive in a ternary fashion. This is illustrated by the example in Table 2, which shows how, in transitive verbs, vowel length distinguishes both unrelated lexical roots and forms within morphological paradigms.

Table 2. A minimal set example of three-level vowel length in Shilluk. Sounds are embedded, click to play.

<table>
<thead>
<tr>
<th></th>
<th>CVC</th>
<th>CVVC</th>
<th>CVVVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>á-thá</td>
<td>á-thut</td>
<td>á-thút</td>
<td>á-thúut</td>
</tr>
</tbody>
</table>
following context is more variable: it consists of one or two syllables, and the following tone target is Low or Mid-level.

The materials include at least six items for each of the four levels of Tone: three with a short vowel, and three with an overlong vowel. For Early High Fall and Late High Fall, the number of items is greater.

2.2. Speakers

The dataset includes data from 9 native speakers of Shilluk (8 male, and 1 female), all of whom are from Tonga, the cluster of villages at the western edge of the Shilluk-speaking area. The second author is among them. Five speakers were recorded in Tonga; the other six were recorded in Juba and Malakal. The number of speakers recorded is actually greater, but several people had difficulty uttering the required utterances fluently in the absence of a realistic communicative setting.

2.3. Elicitation procedure

The data were collected by the authors, both of whom monitored the recordings. Most speakers spoke English, enabling the first author to elicit the utterances. One did not, and a few were not completely fluent, requiring the involvement of the second author in elicitation. However, also for the speakers who did know English, the second author on many occasions helped to clarify which verb inflection was intended. Gestures were used to convey recurrent aspects of the events that the utterances express, e.g. pointing to a person to elicit the 2nd singular. Overall, this procedure enabled us to avoid giving the example of the target verb form.

The data were elicited in two orders, to balance out order of presentation effects. Between one and three realisations were recorded for each item on the elicitation list.

2.4. Data processing

The first two steps in processing were a) chopping up the raw recording files into utterance-sized sound files, and b) segmentation. This dataset consists of 536 tokens. Then the f0 traces were created and checked. The raw traces were visually checked for tracking errors (e.g. octave drops due to non-modal phonation), and the checked traces were then trimmed for spikes, using the trimming algorithm developed by Yi Xu [10]. This is illustrated by the black and green traces in Figure 3.

An important challenge was how to measure alignment: because of the overall shape of the contour (cf. Figure 1), the turning point does not necessarily correspond to a local maximum. Visually, it is usually clear where the high turning point is: it is where the high level f0 starts falling steeply. However, we did not want to determine this turning point on the basis of our own subjective interpretation. Instead, we adopted a fully-automatic procedure. In a first step, the raw trace was stylized, cancelling out f0 changes below a threshold of 1 semitone. We used the stylization feature in Praat [11] to do this. The outcome of this stylization is illustrated by the red trace in Figure 3. Second, in this stylized trace, we extracted the point defining the beginning of the biggest drop in f0, within a domain starting from the beginning of the target word up to the end of the vowel of the syllable that follows the target word. This point was taken to be the high turning point.

The measurements were aggregated over repetitions, ahead of the analyses. Doing so, we went from 536 tokens to 283 types. Below we present the descriptive statistics relating to the high turning point.

3. Results

The mean realizations for the four levels of the factor Tone – averaged across both speakers and items – are summarized in the averaged f0 in Figure 4. These descriptive statistics are based on the traces after trimming but ahead of stylization. These traces are in line with the hypothesis that the four patterns are distinguished primarily through a difference in timing.

3.1. High Turning Point – Alignment

Alignment is expressed by taking the difference between time value of the high turning point and the beginning of the vowel of the target syllable. The descriptive statistics for the alignment of the high turning point are presented in Figure 5.
As explained in Section 2.4, tonal alignment is calculated in a fully-automatic manner. The measures show that the four tone patterns are separated fairly well in terms of tonal alignment. On average, the mean values for alignment are -75 (Low), 16 (Early High Fall), 102 (Late High Fall), and 185 (High).

The results in Figures 4 and 5 do not display the influence of Vowel length. The mean durations for short and overlong vowels in the dataset are 73 and 141 milliseconds, respectively. This is comparable with measurements for V and VVV vowels, respectively, in the same – sentence-medial – context in Dinka. Figure 6 displays the descriptive statistics for the alignment of the high turning point by Tone, with separate panels by Vowel length. The level of Tone that differs the most as a function of Vowel length is the Late High Fall: it follows 77 ms after vowel onset in CVC targets, and 126 ms after vowel onset in CVVVC targets – a difference of 49 ms. While the distance from the start of the vowel varies greatly for the Late High Fall as a function of Vowel length, the distance from the end of the vowel varies little: across CVC and CVVVC targets, the high turning point is near the end of the vowel in this sentence-medial context.

The question, then, is whether any of the pairwise differences between adjacent levels of Tone are too great to be plausibly attributed to declination. This is not case: the pairwise differences between adjacent levels of Tone are all below 1 semitone. The greatest pairwise difference – between Low and Early High Fall on a CVVVC target – is 0.9 semitone.

4. Conclusion

The descriptive statistics are in support with the hypothesis that Shilluk has both an early-aligned and a late-aligned falling contour, and that they are in phonological contrast not only with each other but also with Low and High level tone categories. That is, it is the alignment of the high turning point that distinguishes these four tone patterns following a high plateau, rather than the F0 height of the high turning point.

In our work in progress, we are investigating this dataset using inferential statistics, and we are reflecting on the theoretical and typological relevance of these findings.

5. Acknowledgements

We thank the speakers for their effort in participating in the recording sessions. They are (in addition to the second author): Anesa Nyacam, James Peter Othol, Paulino Okan, Emmanuel Antony, Alexander Otwowo, Francis Boywomo, Amum Obwony, and Elia Otham. We also acknowledge the Summer Institute of Linguistics and the British Council, both in South Sudan, for enabling the first author to carry out fieldwork in there. This research project is supported by the Volkswagen Foundation, as part of the research project Tonal Placement – The Interaction of Qualitative and Quantitative Factors (ToPIQQ), coordinated at University of Cologne by Martine Grice and Anne Hermes.

6. References