Evidence for contrastive alignment in falling contours

Bert Remijsen    Otto Gwado Ayoker

*University of Edinburgh*
The null hypothesis: tonal alignment is not contrastive in contour tones
Null hypothesis: tonal alignment not contrastive in contour tones

“[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.”

[Odden 1995:450]
“[T]here is no possible opposition between two HL or two LH contours where the two tones are synchronized differently within the syllable.”

[Hyman 1988:51]
“If the contours are composed of levels, the existence of two falls implies at least three levels, and then in fact we might expect up to three contours of the same shape.”

[Yip 2002:29]
Null hypothesis: tonal alignment not contrastive in contour tones

This configuration is assumed not to occur:

Figure. Schematic representations of two falling contours distinguished by tonal alignment.
The null hypothesis makes good sense:

- **Production** – F0 changes are implemented slowly, because speed is low at turning points (Xu & Sun 2002)
- **Perception** – when time domain is small, an F0 change has to be greater, to be perceived as a pitch change (‘t Hart, Collier & Cohen 1990)
- **Typology** – contour tones tend to be restricted to environments that offer extra time (Zhang 2001)
But there is some counterevidence
But there is some counterevidence

Dinka

• Evidence of contrastive alignment in falling contours.
• First postulated in Andersen (1987)
But there is some counterevidence

**Dinka**

- **Example:**
  - Early-aligned Fall (allophonic variant of Low toneme)
    - \[ \text{ràa}\text{an } \ddot{\text{a}}-\text{lel} \]
    - person:S AG.S-isolate:3S
    - ‘You isolate a person.’
  - Late-aligned Fall (realisation of Fall toneme)
    - \[ \text{ràa}\text{an } \ddot{\text{a}}-\text{lel} \]
    - person:S AG.S-isolate:PASS
    - ‘A person is being isolated.’
  - ‘He isolates a person.’
    - \[ \text{ràa}\text{an } \ddot{\text{a}}-\text{leel} \]
    - person:S AG.S-isolate:3S
    - ‘A person is being provoked.’
Dinka

• Corroborated by acoustic study
• Early and Late Falls are diverge in tonal alignment (44 ms on average)
• No significant difference in peak height

Figure. Averaged f0 traces on normalised time axis (4 sets, 13 speakers). From Remijsen (2013).

But there is some counterevidence
YoloxóChitl Mixtec

• Evidence of contrastive alignment in rising contours (DiCanio, this workshop)

• Remarkable finding, given that rises take longer to realise than falls (Sundberg 1979, Xu & Sun 2002)
House (1990)

- Evidence from perception experiments, in which tonal alignment is varied across the syllable
- Perception of falling f0 patterns according to House (1990:133ff):

<table>
<thead>
<tr>
<th>F0 contour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmental sequence</td>
</tr>
<tr>
<td>V</td>
</tr>
</tbody>
</table>

But there is some counterevidence
House (1990)

“For the movement contour features Falling [...] and Rising to be optimally perceived, three conditions must be fulfilled”:

1. The fall takes place through a zone of relative spectral stability;
2. The fall sets in 30-50 milliseconds into the vowel;
3. The vowel is at least 100 milliseconds long.

But there is some counterevidence
House (1990)

• “Another area altogether is that of tonal distinctions: what if anything is quantal about them? Candidates for quantal characterizations of tone do not leap immediately to mind.”

[Stevens & Keyser 2010:13]

• House’s hypothesis is in effect a proposal for such a quantal threshold.
The next step
• If we were to make the case that /e/ vs. /ɛ/ is contrastive in a given language, we would want to see:

1. Evidence of substantial and significant differences in F1, F2 in hypothesized (minimal) sets of these two vowels;

2. Evidence that /e/ and /ɛ/ are substantially and significantly different in F1, F2 from /i/ and /a/, respectively.
Similarly in relation to tonal alignment: early- and late-aligned falls could be analysed as /L/ and /H/, respectively (cf. Hyman 2010:203).

Figure. Two falling contours distinguished by tonal alignment.
Confronted with a constellation like the one below, representing the contours as level tones is not an option.

Figure. Schematic representations of four tone categories distinguished by tonal alignment.
This study: report on a production study on such a four-way configuration in phonetic tonal alignment.

Research question: is tonal alignment contrastive in falling contour tone categories?

Null hypothesis: no

Alternative hypothesis: yes
Background on Shilluk
Background on Shilluk

Shilluk

- A Western Nilotic language, within the Nilo-Saharan language family
- Spoken in South Sudan
- At least 200,000 speakers

Figure. Map from Storch (2005), showing the Western Nilotic languages
Background on Shilluk

- Rich in tone: seven distinctive tone patterns
- Rich in quantity: three levels of vowel length
- Tone and vowel length have a high functional load in morphology
Background on Shilluk

- Length, tone have a high functional load in morphology:

<table>
<thead>
<tr>
<th>Verb classes in past tense</th>
<th>Fixed Short</th>
<th>Short with Grade</th>
<th>Long</th>
<th>High Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Fall</td>
<td>Low</td>
<td>Fall</td>
</tr>
<tr>
<td>{lɛŋ}</td>
<td>{lɛŋ}</td>
<td>{cam}</td>
<td>{lɛɛŋ}</td>
<td>{lʊʊɲ}</td>
</tr>
<tr>
<td>‘take’</td>
<td>‘drum’</td>
<td>‘eat’</td>
<td>‘throw’</td>
<td>‘pluck’</td>
</tr>
<tr>
<td>{cam}</td>
<td>{mʌl}</td>
<td>{lɛɛŋ}</td>
<td>{lʊʊɲ}</td>
<td>{lʊʊɲ}</td>
</tr>
<tr>
<td>‘eat’</td>
<td>‘roast’</td>
<td>‘throw’</td>
<td>‘pluck’</td>
<td>‘take turns’</td>
</tr>
<tr>
<td>{mʌl}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘roast’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>{lɛɛŋ}</td>
<td>{lʊʊɲ}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘throw’</td>
<td>‘pluck’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>{lʊʊɲ}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘pluck’</td>
<td>‘take turns’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From Remijsen, Miller-Naudé & Gilley (to appear)
<table>
<thead>
<tr>
<th>Level</th>
<th>Verb Phrase</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>njēn á-ľęŋ kì kęŋ</td>
<td>‘You took money gradually here.’</td>
</tr>
<tr>
<td>High</td>
<td>lōōt á-ľęŋ bûul</td>
<td>‘You used a stick to beat the drum.’</td>
</tr>
<tr>
<td>Mid</td>
<td>mâat á-ľęŋ njēn</td>
<td>‘Somebody used a friend to take money gradually.’</td>
</tr>
<tr>
<td>Low Fall</td>
<td>bûul á-ľęŋ kì kęŋ</td>
<td>‘You beat the drum here.’</td>
</tr>
<tr>
<td>Early High Fall</td>
<td>bûul á-ľęŋ kì kęŋ</td>
<td>‘Somebody has beaten the drum here.’</td>
</tr>
<tr>
<td>Late High Fall</td>
<td>njēn á-ľęŋ pâac</td>
<td>‘Smb. went to the village to take money gradually.’</td>
</tr>
<tr>
<td>Rise</td>
<td>mâat á-ľęŋ njēn</td>
<td>‘You used a friend to take money gradually.’</td>
</tr>
</tbody>
</table>
Background on Shilluk

• An example of three-level vowel length:

CVC

būul á-lút kì kēn

Bol PAST-hit PREP place:S.DEM

‘Somebody has hit Bol here.’

CVVC

kàl á-lúut kì kēn

compound:S PAST-beat PREP place:S.DEM

‘Somebody has carved into the compound here.’

CVVVC

kāc á-lúuut kàl

hoe:S PAST-carveinto:APPL compound:S

‘Somebody has used a hoe to carve into the compound.’
Methods
• Manipulation of the factor Tone, with factor levels:
  - Low
  - Early High Fall
  - Late High Fall
  - High
- Embedded in sentence-medial position:

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Early High Fall</th>
<th>Late High Fall</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRON</td>
<td>P-DEM</td>
<td>S-DEM</td>
<td>S-DEM</td>
<td>S-DEM</td>
</tr>
<tr>
<td>PAST</td>
<td>take:2SG</td>
<td>beat</td>
<td>beat:FUG</td>
<td>beat:APPL.2SG</td>
</tr>
<tr>
<td>PREP</td>
<td>place:S.DEM</td>
<td></td>
<td>village:S</td>
<td>drum:S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gìk-ání á-lèŋ kì kêŋ</td>
<td>gìn-ání á-lèŋ kì kêŋ</td>
<td>gìn-ání á-lèŋ pâac</td>
<td>gìn-ání á-lèŋ bùul</td>
</tr>
</tbody>
</table>

- ‘Somebody has taken it gradually here.’
- ‘Somebody has beaten the it here.’
- ‘Somebody went away to the village to beat it.’
- ‘You have used it to beat the drum.’
• The preceding tonal context is constant (high plateau):

**Low**

given-áni á-len kì kën
PRON:P-DEM PAST-take:2SG PREP place:S.DEM
'Somebody has taken it gradually here.'

**Early High Fall**

given-áni á-len kì kën
PRON:S-DEM PAST-beat PREP place:S.DEM
'Somebody has beaten the it here.'

**Late High Fall**

given-áni á-len pâac
PRON:S-DEM PAST-beat:FUG village:S
'Somebody went away to the village to beat it.'

**High**

given-áni á-len buul
PRON:S-DEM PAST-beat:APPL.2SG drum:S
'You have used it to beat the drum.'
• The following tonal context is Low or Low Fall:

**Low**

*gîk-ání  á-lèŋ  kè  kêŋ*

PRON:P-DEM  PAST-take:2SG  PREP  place:S.DEM

‘Somebody has taken it gradually here.’

**Early High Fall**

*gîn-ání  á-lèŋ  kè  kêŋ*

PRON:S-DEM  PAST-beat  PREP  place:S.DEM

‘Somebody has beaten the it here.’

**Late High Fall**

*gîn-ání  á-lèŋ  pâac*

PRON:S-DEM  PAST-beat:FUG  village:S

‘Somebody went away to the village to beat it.’

**High**

*gîn-ání  á-lèŋ  bûul*

PRON:S-DEM  PAST-beat:APPL.2SG  drum:S

‘You have used it to beat the drum.’
Methods / Design

Low

PRON:P-DEM  PAST-take:2SG  PREP  place:S.DEM
‘Somebody has taken it gradually here.’

Early High Fall

PRON:S-DEM  PAST-beat  PREP  place:S.DEM
‘Somebody has beaten the it here.’

Late High Fall

PRON:S-DEM  PAST-beat:FUG  village:S
‘Somebody went away to the village to beat it.’

High

PRON:S-DEM  PAST-beat:APPL.2SG  drum:S
‘You have used it to beat the drum.’
Methods / Design

- Items (lexical-morphological forms): 3 for Low and High; 6 for Early High Fall and Late High Fall.

- Items have these segmental structures: /lɛŋ, ñɔl, lɛm, ñɛn, łoŋ/

- Characteristics:
  - Invariably CVC
  - Vowel invariably low-mid
  - Consonants invariably /l/ or a nasal

- Forms with long, overlong vowels as fillers (hindsight)

- On average 2 realisations of each form
Methods / Speakers

- The dataset includes recordings with 9 native speakers of Shilluk (8 male, 1 female)
- Dialect is fixed: Tonga, the westernmost cluster of villages in the Shilluk territory (arrow on map).
- Recorded in Tonga, Juba and Malakal

*Figure. Map from Storch (2005).*
Methods / Processing

Step 1: raw traces

Early High Fall

Late High Fall
Methods / Processing

Step 2: spikes trimmed with algorithm of Xu 1999)
• The key measurement: the high turning point that marks the right edge of the high plateau.

• The high turning point is determined automatically, i.e., without subjective decision.

• Stylization in Praat (Boersma & Weenink 2013), with a threshold of 1 semitone.

• Find greatest drop in the stylized trace; assign its initial point to the high turning point.
Methods / Processing

Step 3: stylization

Early High Fall

Late High Fall
Step 4: find high turning point (red dot)
Results
Results / time-normalised plot by Tone

**Figure.** Averaged f0 traces of the target word on a normalised time axis (whole dataset).
Results / tonal alignment by Tone

Figure. Means and standard deviations for tonal alignment of the high turning point by Tone.
Figure. Means and standard deviations for the F0 height of the high turning point, by Tone. Separate panels for raw f0, in hertz (left), and after conversion to ERB scale and z-transformation per speaker (right).
Results / size of F0 drop by Tone

**Figure.** Means and standard deviations for the size of the F0 drop, by Tone. Separate panels for raw f0, in hertz (left), and after conversion to ERB scale and z-transformation per speaker (right).
Results / nucleus and coda duration by Tone

**Figure.** Means and standard deviations for nucleus duration (left) and coda duration (right), by Tone.
## Results / inferential stats

<table>
<thead>
<tr>
<th>Dependent</th>
<th>Factor Tone in ANOVA (within-subjects, df = 3,8)</th>
<th>Pairwise comparisons between adjacent levels of Tone (t-test, Bonferroni)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonal alignment</td>
<td>$F = 36.4, p &lt; .001$</td>
<td>All significant, at $p &lt; .001$</td>
</tr>
<tr>
<td>F0 height (ERB, z-scores)</td>
<td>$F = 0.7, \text{n.s.}$</td>
<td>n.s.</td>
</tr>
<tr>
<td>Size of drop (ERB, z-scores)</td>
<td>$F = 0.4, \text{n.s.}$</td>
<td>n.s.</td>
</tr>
<tr>
<td>Rhyme duration (nucleus+coda)</td>
<td>$F = 0.5, \text{n.s.}$</td>
<td>n.s.</td>
</tr>
<tr>
<td>Dependent</td>
<td>Factor Tone in ANOVA (within-subjects, df = 3,8)</td>
<td>Pairwise comparisons between adjacent levels of Tone (t-test, Bonferroni)</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Tonal alignment</td>
<td>$F = 36.4$, $p &lt; .001$</td>
<td>All significant, at $p &lt; .001$</td>
</tr>
<tr>
<td>F0 height (ERB, z-scores)</td>
<td>$F = 0.7$, n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Size of drop (ERB, z-scores)</td>
<td>$F = 0.4$, n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Rhyme duration (nucleus+coda)</td>
<td>$F = 0.5$, n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
Conclusion, discussion
The hypothesis that tonal alignment can be distinctive in falling contour tones is corroborated.

Conclusion
Discussion / 1. So what about time pressure?
Given that time pressure matters for contour tones (Zhang 2001), how is the Shilluk phenomenon possible?

The contour tones are only found on closed monosyllabic stems. With a coda, the syllable has a minimal duration of about 150 milliseconds, even if the vowel is short. They are not found on function morphemes.

Similar observations apply to Dinka, and to YoloxóChitl Mixtec (DiCanio, this conference)

Generalisation: contrastive alignment in contours is found in environments that offer more space.
Discussion / 2. Early- vs. late-aligned falls in Shilluk vs. Dinka
The early- and late-aligned contours are aligned later in Shilluk than in Dinka:

Table. Mean values for tonal alignment (ms relative to start of vowel) in CVC context in Shilluk and Dinka.

<table>
<thead>
<tr>
<th></th>
<th>Early-aligned Fall</th>
<th>Late-aligned Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shilluk</td>
<td>13</td>
<td>77</td>
</tr>
<tr>
<td>Dinka</td>
<td>-7</td>
<td>33</td>
</tr>
</tbody>
</table>

What does this mean?
Discussion / 2. Early- vs. late-aligned falls in Shilluk vs. Dinka

What does this mean?

• It may be due to a difference in context: the Dinka study used sentence-final context, which shifts alignment earlier (cf. Myers 2003 on Kinyarwanda).

Figure. Peak location, in proportion to duration of the syllable, as a function of lexical tone and sentence context. Data from 1 of the speakers in Myers (2003).
What does this mean?

- It may also be due to the way the High target is measured:
  - as a local F0 maximum in the Dinka study;
  - through stylization in this study.

- Further research is needed to establish the degree of variability due to contextual factors.
Discussion / 3. The question of representation
“If further research uncovers cases where the alignment is contrastive within a language, these might be handled by the use of an alignment feature on the prosodic nodes or on the substantive elements”

[Pierrehumbert & Beckman 1988:159]

In other words, we can enrich the specification on the tonal tier, or on the prosodic constituents with which the tones are linked.
Enriching the tonal tier (Remijsen 2013)

Early High Fall: \((HL)[−\text{late-aligned}]\)

Late High Fall: \((HL)[+\text{late-aligned}]\)

• A feature \([± \text{late-aligned}]\) on the tonal tier, invoked when alignment is contrastive in contour tones.
Enriching prosodic nodes (Prieto, D’Imperio & Gili Fivela 2005: Morén & Zsiga 2006)

Early High Fall:  

\[
\begin{array}{c}
H \\
\mid \\
I \left[ \mu \varepsilon \right] \mu \eta \right]_{\sigma} \\
\mid \\
L \\
\end{array}
\]

Late High Fall:  

\[
\begin{array}{c}
H \\
\mid \\
I \left[ \mu \varepsilon \right] \mu \eta \right]_{\sigma} \\
\mid \\
L \\
\end{array}
\]

- Tones associated with mora edges. In this perspective, the mora is interpreted as a phonetic timing unit.
Comparison of the two approaches

- The feature-based approach is maximally binary
- Associating tones with mora edges allows for more possibilities – e.g. a three-way alignment contrast in contour tones on long vowels:

\[
T = \left[ \sigma C \left[ \mu V \right]_\mu \left[ \mu V \right]_\mu C \right]_\sigma
\]
Discussion / 3. The question of representation

- Feature-based approach is explanatory adequate, given Stevens (1989), House (1990), Stevens & Keyser (2010).

- Interpreting the mora as a timing unit is problematic, if variability is considerable across contexts.

Figure. Peak location, in proportion to duration of the syllable, as a function of lexical tone and sentence context. Data from 1 speaker, in Myers (2003).


Xu & Sun (2002). Maximum speed of pitch change and how it may relate to speech. *JASA* 111, 1399-1413.


Acknowledgements


• The British Council and SIL International, for enabling fieldwork research in South Sudan.

• The Volkswagen Foundation, for research funding through the project “Tonal Placement: Interaction of Qualitative and Quantitative Factors (ToPIQQQ)”, coordinated by Martine Grice and Anne Hermes.
• Number of utterance tokens: 293
• Aggregated over repetitions, ahead of quantitative analyses: 154 types. So there are 8 missing values, given a logical total of 162 (18 items x 9 speakers).
• ANOVAs are based on balanced dataset, with 105 types. So there are 3 missing values, given a logical total of 108 (12 items [3 for each level of Tone] x 9 speakers).