

# TONAL ALIGNMENT IS CONTRASTIVE IN FALLING CONTOURS IN DINKA

BERT REMIJSSEN

*The University of Edinburgh*

This study investigates a contrast in tonal alignment that involves falling contours in Dinka. This contrast calls into question the assumption that tonal alignment cannot distinguish contour tones of the same shape within the syllable domain (e.g. Odden 1995). A qualitative description and a small-scale perception study are followed by a detailed production study. The results indicate that the main correlate of the contrast is indeed tonal alignment: the early-aligned fall sets in during the onset or early in the vowel; the late-aligned fall sets in well into the vowel. The production study also suggests that it is unlikely for more than two patterns of alignment in contour tones of the same shape to be accurately produced and perceived, given various phonetic limitations. The contrast is represented phonologically using a binary feature. This representation is adequate in an explanatory sense, in that the category boundary is in line with the quantal threshold hypothesized in House 1990. The results also corroborate the hypothesis of three-level vowel length in Dinka.\*

*Keywords:* tonal alignment, contour tones, tone features, vowel length, pitch perception, time pressure, Dinka

## 1. INTRODUCTION.

**1.1. MOTIVATION.** Tonal alignment refers to the timing of fundamental frequency ( $f_0$ ) patterns relative to the sequence of speech segments. Many linguists assume that the alignment of tone contours within a syllable is not contrastive (Odden 1995:450, 474, Silverman 1997:479–80, Yip 2002:29). This view is spelled out clearly by Odden: ‘it 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’ (1995:474). This assumption implies that ‘[i]f the contours are composed of levels, the existence of two falls [in a tonal inventory] implies at least three levels’ (Yip 2002:29). In other words, a contrast between two falling contours necessarily involves a difference in the height of the tone targets, rather than in their alignment. The same assumption is implicit in Silverman 1997, which argues that a falling contour resulting from tone sandhi would neutralize with an underlying falling contour within the same range. The common assumption is expressed as a hypothesis in 1.

(1) The alignment of falling  $f_0$  contours within a syllable is not distinctive.

\* I gratefully acknowledge discussions with Johanneke Caspers, Ingo Hertrich, Larry Hyman, Bob Ladd, and Alice Turk. I thank Yi Xu for sharing the algorithm to trim spikes in  $f_0$  traces, and Mike Allerhand and Mateo Obregón for support with statistics. Earlier versions were presented at the ANU Tone Workshop (Canberra, December 4–16, 2011), and at the 20th Manchester Phonology Meeting (Manchester, May 24–26, 2012). The final version has been improved on the basis of an evaluation by associate editor Kie Zuraw, and by two anonymous referees. I thank them for their valuable feedback.

The speech data on which this research is based were recorded from John Penn de Ngong, Peter Garang Nyarjok, Abraham Pach Alier, Job Anyang Aluong, James Maker Riak, and Mary Agotich Buol for Bor South, and from Daniel Akoy Deng, Leek Deng Mawut, Leek Deng Lueth, Akol Kongoor Reech, Abraham Duot de Khueer, Aluel Ajaang Jibol, Abraham Leek Makuei, Emmanuel Deng Jogaak, and Simon Yak Deng for Bor North. The last three also took part in the perception study, as did Elizabeth Achol Ajuet, Peter Malek Ayuel, and David Dau Arop. I gratefully acknowledge their involvement, and in particular John Penn de Ngong, who provided assistance with data elicitation.

All of the data were collected in Juba (South Sudan), in the context of research trips sponsored by SIL Sudan (three trips) and by the British Council (one trip). I thank them, and also the community at the SIL compound in Juba, for hospitality there. Finally, I gratefully acknowledge the Arts & Humanities Research Council (United Kingdom), which funded this research through the project ‘Metre and Melody in Dinka Speech and Song’, as part of the Beyond Text initiative.

This hypothesis is specific to the syllable domain. In other words, it does not rule out instances of contrastive alignment involving a high target followed by a low target, over a domain that extends beyond a single syllable. The latter configuration is well attested—see, for example, Frota 2002 on two intonational tone patterns in European Portuguese. Also beyond the scope of 1 are phenomena whereby two contrasting tone patterns diverge in the alignment of a target, but the overall shape within the syllable is different, as in the contrast between the Fall and High tone categories of Thai (Morén & Zsiga 2006 and further references there). Restricted in this way, 1 is warranted in the face of the available evidence from the world's languages. In his survey of 187 tone languages that have contour tones, Zhang (2001, 2002) makes no mention of phenomena that would challenge it. In fact, Zhang reports that the distribution of contour tones is often restricted to contexts that offer more sonorous duration within the syllable rhyme, such as environments involving a long vowel, a sonorant coda, stress, and utterance-final lengthening. If languages struggle to implement contour tones consistently across contexts, then it is unlikely that they could implement contrastive contour tones of the same shape within a syllable. Several phonetic limitations underlie this interaction between contour tones and sonorous duration within the rhyme. On the production side, there is the time required to implement an  $f_0$  change. The report on this issue by Xu and Sun (2002) suggests that it takes 124 ms on average to realize a fall in  $f_0$  of around four semitones (ST; e.g. 150 to 120 Hz)—a time interval exceeding the duration of a vowel in many environments. On the perception side, one limitation is the glissando threshold, which reflects the hearer's ability to perceive  $f_0$  changes as pitch contours rather than as pitch levels (Rossi 1971, Greenberg & Zee 1979, 't Hart et al. 1990). This threshold is determined by the size and duration of the  $f_0$  change. A second perceptual limitation relates to the difference in timing relative to the segmentals for two otherwise identical sequences of tone targets to be reliably distinguished. Surveying the results of studies bearing on this question, House (2004) finds that the threshold is at or above 50 milliseconds (ms) in most studies. This threshold sensitivity to differences in the timing of a tone target can be wide enough to turn a falling  $f_0$  pattern over the syllable into a rise (D'Imperio & House 1997).

Against this background, this article investigates the phenomenon of contrastive alignment in Dinka, a Nilo-Saharan language. In the context of a descriptive analysis of Dinka phonology, Andersen (1987) postulated a lexical and morphological contrast between two falling contours, which diverge primarily in terms of alignment within the syllable. On the basis of auditory impressions, Andersen (1987:20) describes one as 'fall ... followed by level pitch', and the other as 'level pitch ... followed by fall'. The two contours are identical in terms of Andersen's numerical representation of the excursion size of the pitch fall. In an investigation of a different dialect, Remijsen and Ladd (2008:181) also report that there are two tone categories involving falling contours. Both of these studies present evidence from tone sandhi suggesting that contour tones in Dinka are composed of level targets. As this phenomenon contradicts the widely held generalization in 1, it is worthwhile to examine it in detail. The phenomenon is investigated through a production study in which the amount of sonorous duration is manipulated experimentally (cf. Caspers & van Heuven 1993). If Andersen's hypothesis is correct and the contrast is indeed one of tonal alignment, then the contrasting patterns should diverge primarily in the timing of the targets with the segmental sequence, and this correlate should be stable across manipulations of time pressure.

A corroboration of Andersen's hypothesis would then raise two important issues. First, as noted above, the phonetics of contour tones are constrained by limitations in speech production and speech perception, which may affect their phonological distribu-

tion (Zhang 2001, 2002). If Dinka does show a contrast between two contour tones, it will be worthwhile to find out how this is possible. Dinka is particularly suitable for exploring the impact of these phonetic limitations, because sonorous duration can be controlled effectively through contrastive vowel length.

The second issue on which the corroboration of Andersen's hypothesis would have a bearing is phonological representation. In the autosegmental approach, where tonal phenomena are represented by associating tone targets with metrical units, there are two general strategies for representing contrastive alignment: (i) increasing the level of detail of representation on the tonal tier, and (ii) associating the tone targets with a smaller unit within the prosodic hierarchy. I argue for a feature-based representation on the tonal tier.

The remainder of this introduction provides background on Dinka and its sound system. I then present a description of the alignment contrast between falling contours in Dinka, and provide preliminary support for the hypothesis that there are indeed two falling patterns in the surface phonology. This is followed by the production study (§3), after which I evaluate hypothesis 1 in the face of the evidence from Dinka (§4.1) and consider the above-mentioned issues, namely phonetic limitations (§4.2) and phonological representation (§4.3).

**1.2. BACKGROUND ON DINKA AND ITS SUPRASEGMENTAL PHONOLOGY.** Dinka is a Western Nilotic language within the Nilo-Saharan language family (Gordon 2005). It has approximately two million speakers, primarily in South Sudan. The dialects of Dinka can be divided into four geographic clusters: Padang, Rek, Agar, and Bor (Roettger & Roettger 1989). This article deals primarily with the Bor cluster, which includes—from south to north—the dialects known as Bor (in a narrow sense), Twic, Nyarweng, and Hol. The first three of these are under investigation in the production study in §3. With respect to tone, Twic and Nyarweng do not present any differences between themselves, but the two of them differ from Bor (narrow sense) in a way that is relevant to this study. I therefore refer to Bor (narrow sense) as Bor South (BS), and to Twic and Nyarweng together as Bor North (BN).

A salient characteristic of Dinka phonology is its rich system of suprasegmental contrasts: tone, vowel length, and voice quality are all distinctive, independently of one another (Andersen 1987, Remijsen & Manyang 2009). The minimal set in 2 illustrates the vowel-length contrast. Vowel length distinguishes both unrelated lexical items—for example, the short root {lel} 'isolate' in 2a,b vs. the long root {leel} 'provoke' in 2c—and also related forms within paradigms—for example, 2nd vs. 3rd singular of {lel} 'isolate' in 2a vs. 2b, respectively.<sup>1</sup>

- |                                                                                   |                                                                         |
|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| (2) a. ràaan ā-lèl<br>person DECL.SG-isolate.2SG<br>'You are isolating a person.' | d. r̀̀oor āa-lèl<br>men DECL.PL-isolate.2SG<br>'You are isolating men.' |
| b. ràaan ā-lèel<br>person DECL.SG-isolate.3SG<br>'He is isolating a person.'      | e. r̀̀oor āa-lèel<br>men DECL.PL-isolate.3SG<br>'He is isolating men.'  |
| c. ràaan ā-lèeel<br>person DECL.SG-provoke.3SG<br>'He is provoking a person.'     | f. r̀̀oor āa-lèeel<br>men DECL.PL-provoke.3SG<br>'He is provoking men.' |

<sup>1</sup> Abbreviations in the glosses are as follows: DECL: declarative, PASS: passive, PL: plural, PREP: preposition, SG: singular, U: uncountable. Audio files for examples 2, 6, 7, 9, and Fig. 11 below are linked to the .pdf and .html versions of this article, and can also be accessed at <http://muse.jhu.edu/journals/language/v089/89.2.remijsen01.html>.

As seen from 2a–c, the combination of lexical and morphological length combines into a three-level vowel-length contrast in content morphemes. The hypothesis of three-level vowel length was first proposed for Dinka by Andersen (1987, 1993) and has been corroborated on the basis of phonetic evidence (Remijsen & Gilley 2008). The three levels of vowel length—short, long, and overlong—can be represented through one, two, and three moras, respectively, associated with the syllable nucleus (Flack 2007, Remijsen & Gilley 2008:320). This is illustrated schematically in 3.



All verb roots consist of a single closed syllable, a pattern that predominates throughout the lexicon (Andersen 1987, 1993). Short verb roots like {lel} ‘isolate’ have a phonologically short stem vowel (V) in some inflections, and a phonologically long stem vowel (VV) elsewhere—for example, /ləl/ ‘isolate.2SG’ vs. /ləél/ ‘isolate.3SG’. Long roots such as {leel} ‘provoke’ have a phonologically long stem vowel in some inflections, and a phonologically overlong stem vowel (VVV) elsewhere—for example, /léel/ ‘provoke.2SG’ vs. /ləéel/ ‘provoke.3SG’. In function morphemes, vowel length is distinctive in a binary fashion, that is, short vs. long. This contrast is illustrated in 2 by the declarative prefix, which is /ā-/ when the preceding topic noun is grammatically singular (1a–c), and /āā-/ when the preceding noun is grammatically plural (1d–f).

Like vowel length, tone is distinctive both between lexical items and within paradigms. Depending on the dialect, the inventory of underlying tone categories or tonemes consists of three or four patterns, associated at the level of the syllable. The Bor dialects all have four tonemes: Low, High, Mid, and Fall. Figure 1 illustrates the realizations of these patterns.<sup>2</sup>

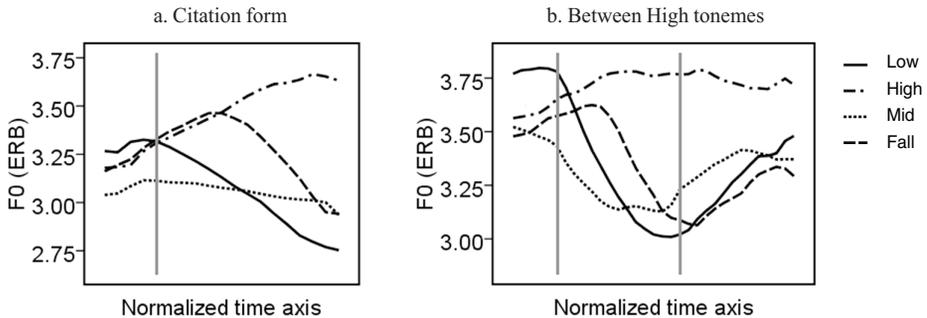


FIGURE 1. Averaged interpolated  $f_0$  traces on a normalized time axis, showing the realization of the four tonemes—Low, High, Fall, and Mid—in /nòoon/ ‘grass.sg’, /lɔ̄ɔ̄om/ ‘rib.sg’, /ŋâap/ ‘sycamore.sg’, and /lāaac/ ‘urine.u’, respectively. Each trace is averaged across realizations by three speakers of Bor North. Panel a: citation forms. Panel b: same target words in a frame sentence, between High tonemes. The vertical line in (a) and that on the left in (b) mark the start of the target word. The second vertical line in (b) marks the end of target word.

There are two phonological categories of voice quality: modal vs. breathy. Like tone and vowel length, voice quality distinguishes unrelated lexical items—for example, /ròoor/ ‘forest’ vs. /ròoor/ ‘men’—and also forms within morphological paradigms, as in

<sup>2</sup>  $F_0$  is expressed using the equivalent rectangular bandwidth (ERB). This scale provides a more accurate representation of human pitch perception than the hertz scale (Hermes & van Gestel 1991, Nolan 2003).



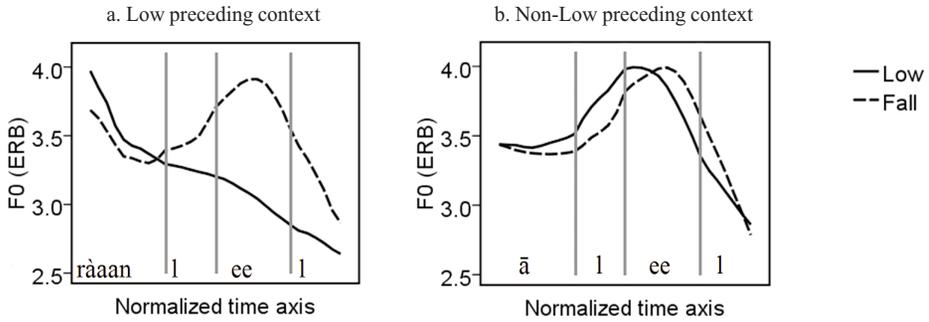


FIGURE 2. Averaged f0 traces on a normalized time axis, illustrating Low vs. Fall tonemes. Each trace is averaged across realizations by four speakers of Bor South. Panel a: last two syllables in 5a,b. Panel b: last two syllables in 6a,b.

- (6) a. rāaan ā-lèel  
 person DECL.SG-isolate.3SG  
 ‘He is isolating a person.’  
 b. rāaan ā-lèel  
 person DECL.SG-provoke.PASS  
 ‘A person is being provoked.’

Averaged traces are presented in Figure 2b. A comparison of the dashed line traces in Figs. 2a vs. 2b shows that the realization of the Fall toneme does not differ noticeably as a function of the preceding tonal context. But the realization of the Low toneme (solid lines) is markedly different in the two contexts. Preceded by another Low toneme, the Low toneme is realized through a low and level f0 pattern (Fig. 2a). Hereafter this allophonic variant is referred to as  $Low^{Level}$ . Preceded by the Mid toneme, however, the realization of the Low toneme involves an f0 peak early in the vowel (Fig. 2b), followed by a falling contour. Hereafter this allophonic variant is referred to as  $Low^{Fall}$ . The difference in realization between the  $Low^{Fall}$  and the Fall tonemes appears to be primarily a matter of the alignment of the f0 peak that defines the starting point of a fall in f0 within the rhyme. This peak is located near the beginning of the vowel in the case of the  $Low^{Fall}$ . In the case of the Fall, this peak is located well into the vowel. I do not perceive any difference in phonetic voice quality between the  $Low^{Fall}$  and the Fall.

These observations on the realization of the  $Low^{Fall}$  and the Fall in Bor South are in line with Andersen’s (1987:20) study on the Agar dialect. As noted above, he summarizes his syllable-level pitch impressions as ‘fall ... followed by level pitch’ in the case of what I have referred to as the  $Low^{Fall}$ , and ‘level pitch ... followed by fall’ in the case of the Fall. Similarly, Andersen does not report a difference in perceived voice quality between the two falling contours.

The distributions of  $Low^{Level}$  and  $Low^{Fall}$  do not overlap. The  $Low^{Level}$  realization is found following another instance of the Low toneme. The  $Low^{Fall}$  is found elsewhere: following the Mid toneme (Fig. 2b), the High toneme (Fig. 1b), and also in the citation form (Fig. 1a). The fact that the  $Low^{Fall}$  is found in the citation form and following the Mid toneme provides insight, because in these contexts, the high target at the beginning of the vowel cannot be attributed to a high end target of a preceding syllable. That is, the Mid toneme does not involve a high end target in the Bor dialects in any other context (cf. Figs. 1a,b).

In most of the dialects of Dinka that have been investigated so far, the Fall is also found on short vowels. If a dialect allows this configuration, then one can expect to find

minimal contrasts of Low<sup>Fall</sup> vs. Fall on syllables with a short vowel, as in the example from Bor North in 7. Comparable contrasts have also been reported for Agar Dinka (Andersen 1987:18) and for Luanyjang Dinka (Remijsen & Manyang 2009:120).

- (7) a. rāaan ā-lèl  
 person DECL.SG-isolate.2SG  
 ‘You are isolating a person.’  
 b. rāaan ā-lèl  
 person DECL.SG-isolate.PASS  
 ‘A person is being isolated.’

In conclusion, the Low<sup>Fall</sup> and Fall tone patterns appear to differ primarily in the alignment within the syllable of a peak that defines the beginning of a falling contour (see Fig. 2b). This observation is in line with the description in Andersen 1987. However, such a state of affairs is at odds with the hypothesis (1) that the timing of f0 changes within the syllable is not contrastive. This hypothesis predicts that, if the Low toneme is realized as a falling contour in certain contexts, it should neutralize with the underlying Fall (cf. Silverman 1997). The following two subsections present evidence as to whether the Low<sup>Fall</sup> and Fall neutralize, from a perception experiment (§2.2) and from a tone sandhi process (§2.3).

**2.2. PERCEPTION EXPERIMENT.** To the best of my knowledge, native speakers of Dinka have no difficulty in disambiguating forms in which the lexical and/or grammatical meaning rests solely on the contrast between the Low<sup>Fall</sup> and the Fall. The experiment reported here constitutes a small-scale attempt to verify this. The subjects are presented with a stimulus whose interpretation hinges crucially on whether the verb has the Low<sup>Fall</sup> or the Fall. After hearing the stimulus utterance, the subjects select between the English translations appropriate for the two interpretations (forced-choice design). For example, one stimulus is a recording of the sentence in 6a, and the subjects respond by selecting either the translation in 6a or the one in 6b. There are ten stimuli for tone, made up of five minimal sets for the contrast between Low<sup>Fall</sup> and Fall, including the sets in 6 and 7. Three minimal sets have a short vowel in the stem syllable, and two have a long vowel. The stimuli come from three different speakers. There is also a control set involving a contrast in vowel height,<sup>4</sup> as an independent criterion on the basis of which to exclude responses from participants who would have difficulties with the task.

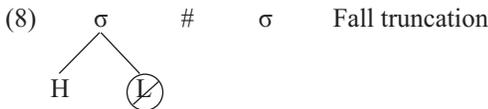
The design was implemented in Praat (Boersma & Weenink 2010) using the Multiple Forced Choice tool for designing and running listening experiments. To gauge the level of consistency in the responses, each stimulus recurs five times in the experiment. The stimuli were presented in five randomized blocks consisting of twelve stimuli each, with each stimulus occurring only once in each block. The stimuli were presented over loudspeakers, and the responses were recorded by clicking the selected English translation. Because some of the subjects were not accustomed to using a computer, the subjects pointed out their selection, and I recorded their choice by clicking the sentence they had pointed at. Six native speakers of Dinka participated as subjects. Three of them are speakers of Bor North (two Twic, one Nyarweng); the other three speak dialects from the Rek cluster (two Twic,<sup>5</sup> one Malual). All of the dialect varieties represented among the participants allow for the Fall toneme on short vowels.

<sup>4</sup> The control set involved the verb forms /ā-wèel/ ‘DECL.SG-turn.over.1SG’ and /ā-wèel/ ‘DECL.SG-turn.over.3SG’.

<sup>5</sup> The Twic dialect in the Rek cluster is different from the Twic dialect in the Bor cluster.

None of the participants had difficulties with the task: each subject invariably chose the correct translation for the control-set items. These are not considered further. The responses on tone stimuli number 300 in total (6 subjects \* 10 stimuli \* 5 repetitions). The correct classification score on these stimuli is 92% (276 correct out of 300). Separated by the dialect region of the participants, the correct classification result is higher for the Rek subjects (95.3% correct) than for the Bor North subjects (88.6% correct). Across all subjects, correct classification was better on stimuli with a short vowel (93.9% correct) than on the stimuli with a long vowel (89.2%). It is clear that the participants were able to infer the lexical and morphological meaning in minimal pairs for tone (Low<sup>Fall</sup> vs. Fall) at a level well above chance. In addition, it is noteworthy that the Rek speakers were also successful at the task, even though the stimuli were drawn from a Bor dialect. This confirms that the contrast between falling contours is representative beyond the Bor dialects (cf. Andersen 1987, 1993, Remijsen & Manyang 2009).

**2.3. TONE SANDHI.** The hypothesis in 1 predicts that when the Dinka Low toneme is realized as a falling contour, it should no longer be phonologically distinct from the Fall toneme (cf. Silverman 1997). Concretely, this hypothesis entails that underlying stem forms like /lèel/ ‘isolate.3SG’ and /lêel/ ‘provoke.PASS’ are indistinguishable as falls in the surface phonology, in, among others, forms involving the Mid-toned declarative prefix. That is, underlying /lèel/ and /lêel/ would both yield surface-phonological /ā-lêel/. The Bor South dialect presents a tone sandhi process that makes it possible to evaluate this prediction. The argument depends on the assumption that phonological processes that apply within words take effect before phonological processes between words (Kiparsky 1985). In the Bor South dialect, the Fall toneme is replaced by the High toneme whenever it is not immediately followed by an utterance boundary. Similar sandhi processes are reported for other dialects (Andersen 1987, Remijsen & Ladd 2008). This process is formulated in autosegmental terms in 8, where the Fall is represented as High-Low. The second component of a High-Low sequence associated with a single syllable is deleted if the tone-bearing unit is followed by another word within the same utterance.



Consider the application of this rule on the verb forms familiar from 6. In 9, these verb forms are followed by an adverb, so that the contextual condition for the rule in 8 is met. If the prefix /ā-/ has the effect of turning a Low toneme into a Fall toneme in the context of the word-internal phonology, then the verb stem in 9a has the Fall toneme at the postlexical stage of derivation, the point where Fall truncation applies.

- (9) a. ràaan ā-lèel                      é-tēnē  
       person DECL.SG-isolate.3SG PREP-here  
       ‘He is isolating a person here.’
- b. ràaan ā-lêel                      é-tēnē  
       person DECL.SG-provoke.PASS PREP-here  
       ‘A person is being provoked here.’

The Low<sup>Fall</sup> pattern is conditioned not only by a word-internal non-Low target, as in 6a, but also by a non-Low target separated by a word boundary, as in Fig. 1b. Given this, any conditioning process producing the Low<sup>Fall</sup> would also have to apply postlexically. By implication, the Low<sup>Fall</sup> may also be conditioned after Fall truncation (8) has applied. I assume that any postlexical conditioning of the Low<sup>Fall</sup> would be vacuous when the conditioning context is already met word-internally.

Figure 3 presents averaged traces based on data from four speakers. In the case of the underlying Fall (9b), the graph shows  $f_0$  rising throughout the rhyme, indicating that the underlying Fall (HL) has turned into High, in line with Fall truncation. This contrasts with the falling  $f_0$  contour in utterance-final position that was seen in Fig. 2b (dashed line).

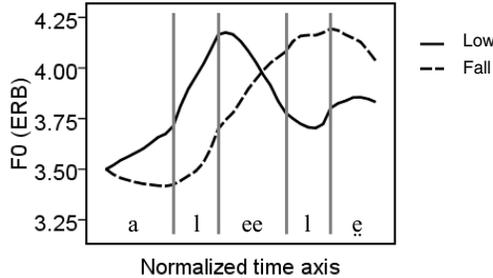


FIGURE 3. Averaged  $f_0$  traces on a normalized time axis, illustrating the realization of syllables underlyingly specified for Low and Fall tonemes, averaged across realizations by four speakers of Bor South. The traces represent the middle section of the sentences in 9a and 9b, respectively.

In contrast, the stem with underlying specification for the Low toneme (9a) shows no evidence of Fall truncation (Fig. 3, solid line). Its averaged trace is very similar to the one in Fig. 2b: there is a falling contour over the stem syllable of the verb, and the beginning point of this contour is close to the beginning of the vowel. On the assumption that the phonological processes are ordered as outlined here, Fall truncation does not affect the  $\text{Low}^{\text{Fall}}$ , suggesting that  $\text{Low}^{\text{Fall}}$  is distinct from the Fall in phonological terms.

**2.4. SUMMARY.** The findings from the perception experiment and from the tone sandhi process are in line with the observations reported in §2.1 and in Andersen 1987, that is, that Dinka has two falling contour tones in its surface phonology. On this basis, I present the hypothesis in 10.

- (10) Tonal alignment distinguishes the  $\text{Low}^{\text{Fall}}$  (early-aligned in syllable) from the Fall (late-aligned in syllable) in the surface phonology of Dinka.

This hypothesis is not compatible with the hypothesis in 1. In the following section I report on a production study in which the phonetic realizations of the  $\text{Low}^{\text{Fall}}$  and the Fall are studied in detail. What is at issue is whether the contrast under investigation is primarily one of tonal alignment. The hypothesis in 1 would be corroborated if the two falling contours diverge primarily in terms of another correlate, such as  $f_0$  height (Yip 2002:29). If, in contrast, the hypothesis in 10 is corroborated, then it will be worthwhile to find out how a contrast in tonal alignment can be implemented phonetically, and to consider how such a contrast can best be represented.

**3. PRODUCTION STUDY.** The goal of this production study is to gain insight into the phonetic realization of the  $\text{Low}^{\text{Fall}}$  and Fall in Dinka. The design hinges on the manipulation of time pressure, which enables the researcher ‘to separate the properties of pitch movements into categories of greater or lesser importance’ (Caspers & van Heuven 1993:162). Factors that are often manipulated in this context include rate of speech (Caspers & van Heuven 1993, Xu 1998, 2001), vowel length and the composition of the syllable (Caspers & van Heuven 1993, Xu 1998, Schepman et al. 2006), and word length (Myers 2003, Yu 2008). In the context of the current study, any parameters that saliently and consistently distinguish  $\text{Low}^{\text{Fall}}$  and Fall from one another across manipu-

lations of time pressure have a bearing on the phonological nature of the two contours, and by extension, on the evaluation of the competing hypotheses in 1 and 10.

### 3.1. METHODS.

**MATERIALS.** The phonological patterns of tone and vowel length in the target words are illustrated in 11. All of the target words involve the Mid-toned declarative prefix /ā(a)-/, which conditions the Low<sup>Fall</sup> realization of the Low toneme, thereby setting up a minimal contrast between Low<sup>Fall</sup> and Fall. Time pressure in the two syllables nearest to the falling f0 change was manipulated independently, by controlling vowel length in the prefix syllable and in the stem syllable. It is relatively straightforward to manipulate time pressure in this way in Dinka, because vowel length is determined by lexical and morphological meaning (see §1.2).

(11) ā-lèl	ā-lèel	ā-lèeel
DECL.SG-isolate.2SG	DECL.SG-isolate.3SG	DECL.SG-provoke.3SG
āa-lèl	āa-lèel	āa-lèeel
DECL.PL-isolate.2SG	DECL.PL-isolate.3SG	DECL.PL-provoke.3SG
ā-lèl*	ā-lèel	
DECL.SG-isolate.PASS	DECL.SG-provoke.PASS	
āa-lèl*	āa-lèel	
DECL.PL-isolate.PASS	DECL.PL-provoke.PASS	

The asterisked forms in 11, that is, the combination of the Fall with a short stem vowel, are found only in Bor North, and as a result they were recorded only from speakers of that dialect. In Bor South, the same inflection is marked instead by the Mid toneme, as in 4b. These Mid-toned forms were recorded from the speakers of Bor South, but they are not included in the analysis of the falling contours reported in §3.2. Not included at all in the design is the combination of the Fall toneme with an overlong stem vowel. This combination is part of the Dinka phonological system, but it was left out for practical reasons.<sup>6</sup> This is not a problem, because time pressure is smallest in forms with an overlong vowel.

In 11, the design of the materials is illustrated on the segmental sequence /lèl-leel-leeel/. The same phonological patterns were also collected for three other sets: /maŋ-maaŋ-maaan/, /ŋop-ŋoop-ŋooot/, and /wel-weel-weeel/. In each case, the onset is a sonorant, and the segments are matched for phonetic type within each set. The target verb forms are embedded in final position in a two-word sentence, preceded by a noun that is monosyllabic and Low-toned, as in 2, 6, and 7.

**SPEAKERS.** Thirteen speakers of Dinka (eleven men, two women) were recorded. Seven were speakers of Bor North (six male, one female) and six of Bor South (five male, one female). All speakers were born in Dinka villages, but spent one to two decades in neighboring countries as refugees. Through a questionnaire, I made sure that all of the participating speakers had spent their time away living with people of the same dialect. The speakers also speak English, plus one or more other languages, such as Arabic, Swahili, and Acholi.

**PROCEDURE.** The data set is built up around eight lexical verb roots, listed in Table 1. That is, each of the four segmentally controlled sets involves two roots—for example,

<sup>6</sup> The combination of the Fall on an overlong vowel is found in, among other places, long verbs inflected for spatial deixis. In pilot data collection, it became clear that these inflections are not available for all verbs, and that they are difficult to elicit in a controlled recording session. For this reason this combination was not included in the design.

{lel} ‘isolate’ and {leel} ‘provoke’ in the case of the /lel-leel-leeel/ set. Inflections of the same verb root were elicited together. The eight root-based blocks were ordered so as to avoid lexical confusion between segmentally similar roots. Table 1 presents the list of stem forms that were elicited in forward order. This order was reversed for half of the speakers, to counterbalance for order-of-presentation effects. Within each block, the order of elicitation was fixed: 1st/3rd singular came first, then 2nd singular, then passive. For each stem inflection, forms preceded by a singular topic noun were elicited before forms with a plural topic noun. Two or three realizations were recorded of each form. The data were elicited using English by the author, initially with the help of a Dinka assistant, John Penn de Ngong, to help ensure that the intended form was recorded. The recordings were made in Juba, South Sudan, using a directional headset-mounted microphone and a solid-state recorder.

VERB ROOT		INFLECTIONS		
		(each elicited with prefixes /ā-/ ‘DECL.SG’ and /āa-/ ‘DECL.PL’)		
{leel}	‘provoke’	3SG: -lèeel		PASS: -lèel
{wel}	‘turn over’	3SG: -wèel	2SG: -wèl	PASS: -wèl (BN)/-wèl (BS)
{maan}	‘hate’	1SG: -māaan		PASS: -māan
{ŋop}	‘take a gulp’	3SG: -ŋòop	2SG: -ŋòp	PASS: -ŋòp (BN)/-ŋòp (BS)
{maŋ}	‘slap’	1SG: -māaŋ	2SG: -māŋ	PASS: -māŋ (BN)/-māŋ (BS)
{ŋoot}	‘spit’	3SG: -ŋòoot		PASS: -ŋòot
{lel}	‘isolate’	3SG: -lèel	2SG: -lèl	PASS: -lèl (BN)/-lèl (BS)
{weel}	‘mess up’	3SG: -wèeel		PASS: -wèel

TABLE 1. The inflections elicited for each verb root, in forward order.

**DATA PROCESSING AND ANALYSIS.** The recordings were processed and analyzed using Praat (Boersma & Weenink 2010). The target word within each sentence was segmented based on information in the waveform and in the spectrogram. In three of the four sets, the onset is a nasal or /l/. These consonants lend themselves well to the study of tonal phenomena, because they carry fundamental frequency without perturbing it upward or downward. It is not uncommon, however, for the  $f_0$  trace to show a spike at the boundary between such a consonant and a vowel. These spikes were corrected using the trimming algorithm described in Xu 1999. The threshold value above which spikes are trimmed was set at 1 Hz. The effect of this correction can be seen in Figure 4. In both of the examples, the spikes in the raw trace (gray) would have shifted the timing measurement of the  $f_0$  peak toward the segment boundary. The corrected trace is overlaid in black. Raw and trimmed traces were compared visually for each target word token, to examine the output of the trimming algorithm and to correct any tracking errors. Figure 4 also illustrates that, for the Low<sup>Fall</sup> and the Fall alike, there is a salient  $f_0$  peak in the context of a preceding Mid toneme. As a result, the height and timing properties of the peak are easy to measure.

The combination of thirteen speakers each producing ten forms for each of four segmental sets yields an expected total of 520 cells. However, fifteen of these are empty. These gaps are due either to items being accidentally skipped during elicitation, or to my failure to notice during elicitation that a speaker uttered a sentence other than the intended one.

The following measures were extracted for each token: (i) the duration of the stem vowel; (ii) the duration of the prefix vowel; (iii) the difference in time between the  $f_0$  peak within the stem syllable and the beginning of the stem vowel, the end of the stem vowel, and the beginning of the prefix vowel; (iv) the  $f_0$  value at this maximum; and (v) the excursion size of the  $f_0$  fall. The values for these measurements were averaged

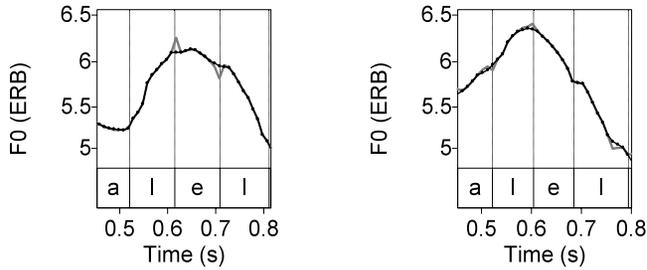


FIGURE 4. Raw (gray) and trimmed (black) f0 traces for the target words in two utterances by a female speaker of Bor North (Twic). Left panel: /āa-l̄l̄/; right panel: /ā-l̄l̄/.

across any repetitions of the same sentence by the same speaker, before running the statistics.

All of these measurements are investigated through descriptive statistics, and those that may constitute correlates of the contrast between  $\text{Low}^{\text{Fall}}$  and Fall are examined further through linear mixed-effects modeling and likelihood ratio tests, as implemented in R (R Development Core Team 2010). The design consists of the random factors Speaker (thirteen) and Set (four), and the fixed factors Tone ( $\text{Low}^{\text{Fall}}$ , Fall), Stem length (V, VV, VVV), Prefix length (V, VV), Dialect (BS, BN), and Coda type (Sonorant, Plosive). Stem forms that have the Mid toneme (i.e. passive forms of short verbs in Bor South) are not included in these analyses. Test values and probabilities are reported for all main effects, and also for any interactions that significantly improve the model in a likelihood ratio test (Bates 2005).

### 3.2. RESULTS.

**VOWEL DURATION.** The measurements for vowel duration—in particular regarding the short and long levels of vowel length in both prefix and stem syllables—are ancillary to the main focus of the production study, in that they demonstrate the extent to which sonorous duration has been varied experimentally. These measurements are also of interest in themselves, because three-level vowel length is highly uncommon among the world's languages (Odden 2011).

The descriptive statistics for vowel length in the stem syllable are presented in Table 2. Long vowels are 59% (40 ms) longer than short vowels, and overlong vowels are about 77% (80 ms) longer than long vowels. A separate set of statistics is provided for the forms that have the  $\text{Low}^{\text{Fall}}$  on the stem syllable, because these represent a data set for three-level vowel length in Dinka in which tone is controlled. In the prefix, where the vowel-length distinction is binary rather than ternary, the average durations were 76.2 ms for the short vowel /ā-/ and 137.9 ms for long /āa-/, that is, a difference of about 80% (60 ms).

STEM LENGTH	$\text{LOW}^{\text{FALL}}$ AND FALL	$\text{LOW}^{\text{FALL}}$ SUBSET
Short (V)	70.7 (12.1)	68.9 (11.7)
Long (VV)	112.8 (16.2)	109.6 (15.7)
Overlong (VVV)	194.4 (30.1)	194.4 (30.1)

TABLE 2. Means and standard deviations (in parentheses) for the duration of the stem vowel (in ms), for the whole data set, and for a subset consisting of those forms that have the  $\text{Low}^{\text{Fall}}$  on the stem syllable.

Figure 5a shows the descriptive statistics for the whole data set as a graph. Comparable results for the Luanyjang dialect, from Remijsen & Gilley 2008, are presented alongside (Figure 5b). In both data sets, the target words are in utterance-final position.

The pattern is similar in the two dialects. The separation between levels of vowel length is clearer in the Bor dialects, where there is no overlap between the distributions at the level of one standard deviation. Factors that are likely to give rise to variation within levels of vowel length include between-speaker variation in rate of speech and intrinsic vowel duration.

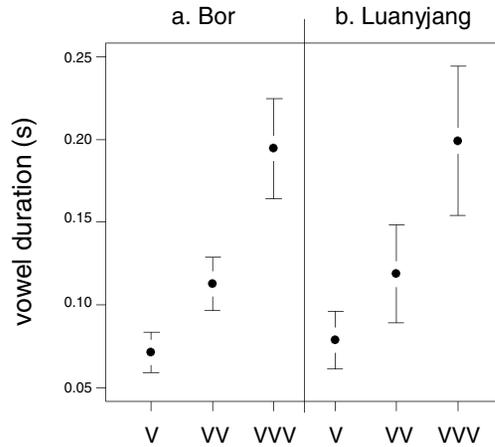


FIGURE 5. Means and standard deviations for vowel duration in the stem syllable, as a function of Stem length (V, VV, VVV). Panel a: this study (thirteen speakers of Bor Dinka). Panel b: comparable data from Remijsen & Gilley 2008 on Luanyjang Dinka.

Table 3 reports the results of the linear mixed-effects model with vowel duration in the stem syllable as the dependent. These results confirm that there is a substantial difference in the duration of the stem vowel as a function of Stem length. The comparisons between levels of Stem length are all significant, characterized by double-digit *t*-values. There is no significant effect for Dialect. This indicates that the vowel-length distinction is realized similarly in Bor North and Bor South. Prefix length and Coda type also do not register a significant effect. There is, however, a significant effect of Tone. Vowel durations are 5–6 ms shorter when the stem is specified for the Low<sup>Fall</sup> than when it carries the Fall. The mean values for Low<sup>Fall</sup> and Fall are 68.9 and 74.1 ms, respectively, when the stem vowel is short, and 109.6 and 115.8 when the vowel is long.<sup>7</sup> In summary, the three-level vowel-length distinction is saliently realized in Dinka Bor. Specifically for the purposes of the current study, it is clear that time pressure was manipulated effectively by controlling Stem length: differences between adjacent levels of vowel length in the prefix and in the stem range from 60 to 80% relative to the lower of the two levels. Vowel duration is also significantly affected by Tone, although this difference is proportionally small: the vowel is between 5 and 7.5% longer if the syllable has the Fall toneme than if it has the Low<sup>Fall</sup>.

**ALIGNMENT OF THE F<sub>0</sub> PEAK.** Figure 6 displays averaged f<sub>0</sub> traces over the stem syllable, as a function of Stem length, Tone, and Coda type. We first consider the alignment of the f<sub>0</sub> peak, then the height of the f<sub>0</sub> peak, and then the excursion size of the fall that follows the peak.

The graphs in Figure 6 suggest that the peak is aligned close to the beginning of the stem vowel in the case of the Low<sup>Fall</sup>, and well into the stem vowel in the case of the

<sup>7</sup> For the sake of comparison, the mean duration for the Mid toneme on a short vowel in Bor South is 73.9 ms.

FACTOR AND LEVELS	<i>t</i> -VALUE	<i>p</i>
Stem length: V vs. VV	24.70	< 0.001
	63.00	< 0.001
	41.90	< 0.001
Prefix length: V- vs. VV-	-1.50	0.126
Coda type: sonorant vs. plosive	-1.06	0.288
Tone: Low <sup>Fall</sup> vs. Fall	3.31	0.001
Dialect: BS vs. BN	0.10	0.909

TABLE 3. Summary of the results of a linear mixed-effects model with the duration of the vowel of the stem syllable as the dependent.

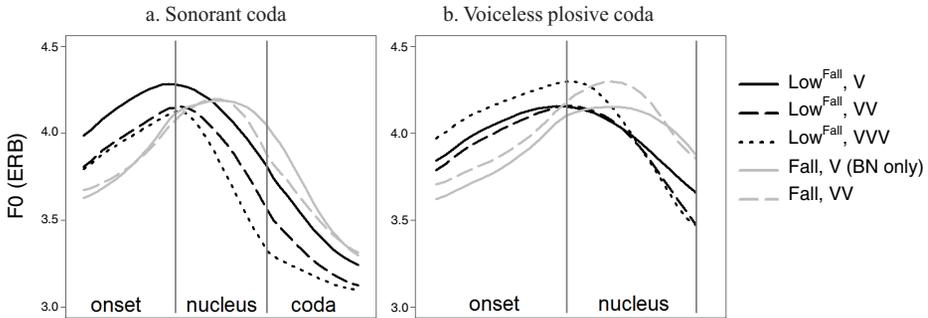


FIGURE 6. Averaged f0 traces on a normalized time axis, by Stem length and Tone. Separate panels for sets ending in a sonorant coda (a) vs. sets ending in a plosive (b).

Fall. To begin with, we need to determine the segmental landmark relative to which the alignment of these peaks can best be expressed. This question can be answered by examining alignment across manipulations of vowel length: the most appropriate landmark is the one relative to which peak alignment varies least. As seen from Table 4, if we use the beginning of the stem vowel as a point of reference, then the peak of the Low<sup>Fall</sup> is aligned on average 9.6 ms later on a long vowel than on a short vowel, and 13 ms later on an overlong vowel than on a short vowel. In contrast, if peak alignment for the Low<sup>Fall</sup> is expressed relative to the end of the vowel, the variations in peak alignment are much greater: 31.3 ms difference between short and long vowels, and as much as 111.4 ms difference between short and overlong vowels. Table 4 also presents comparable descriptive statistics for the Fall. Peak alignment for the Fall differs by 13.1 ms between short and long vowels when it is expressed relative to the beginning of the vowel. Expressed relative to the end of the vowel, the difference is 28.6 ms, that is, more than double. In summary, for the Low<sup>Fall</sup> and the Fall alike, peak alignment can be expressed more appropriately relative to the beginning of the stem vowel than to its end, given that alignment varies less across contexts relative to the former landmark than relative to the latter.

TONE	STEM LENGTH	PEAK ALIGNMENT, EXPRESSED RELATIVE TO:	
		START OF STEM VOWEL	END OF STEM VOWEL
Low <sup>Fall</sup>	CVC	-5.6	-74.4
	CVVC	4.0	-105.7
	CVVVC	8.6	-185.8
Fall	CVC	32.6	-41.5
	CVVC	45.7	-70.1

TABLE 4. Means for f0 peak alignment (in ms) calculated relative to the beginning and the end of the stem vowel, as a function of Tone and Stem length.

Particularly in relation to the Low<sup>Fall</sup>, it is also worthwhile to consider the phonetic alignment of the peak relative to the prefix vowel. Given that the Mid associated with the prefix is one of the contexts in which the Low<sup>Fall</sup> is found, it could be that the peak is aligned in a constant timing relation with this prefix vowel. As seen from Table 5, this is not the case: the Low<sup>Fall</sup> reaches its peak about 59 ms later following a long prefix vowel than after a short prefix vowel. The same difference as a function of Prefix length is found for the Fall. In both cases, this distance between the peak and the beginning of the prefix vowel is instead greatly affected by vowel length: the duration of long vowels is 62 ms greater than that of short vowels.

tone	Prefix Length	Peak Alignment Relative to Start of Prefix Vowel
Low <sup>Fall</sup>	V-	163.5
	VV-	222.3
Fall	V-	199.3
	VV-	258.0

TABLE 5. Means for f0 peak alignment (in ms) calculated relative to the beginning of the prefix vowel, as a function of Tone and Prefix length.

Given these comparisons, further descriptive and inferential statistics on peak alignment will be expressed relative to the beginning of the vowel of the stem syllable. Across the data set, the mean values for peak alignment for the Low<sup>Fall</sup> and the Fall are 2.3 ms and 41.0 ms into the vowel, respectively, in each case with a standard deviation of 15 ms. Means by levels of Dialect, Stem length, and Prefix length are presented in Table 6.

Dialect	Tone	Stem Length	Prefix Length	Prefix Length	Both
			V-	VV-	
Bor North	Low <sup>Fall</sup>	V	-5.6	-12.1	-8.8
		VV	-2.3	-1.8	-2.1
		VVV	7.9	0.0	3.6
	Fall	V	32.9	32.3	32.6
		VV	45.0	41.9	43.5
Bor South	Low <sup>Fall</sup>	V	1.2	-4.7	-1.8
		VV	12.1	8.8	10.5
		VVV	16.1	13.3	14.7
	Fall	VV	48.6	48.9	48.7

TABLE 6. Means (in ms) for the alignment of the peak that defines the beginning of the falling contour, by Dialect, Tone, Stem length, and Prefix length.

Figure 6 and Table 6 both show a difference in peak alignment between Low<sup>Fall</sup> and Fall. This difference is displayed more clearly in Figure 7. This graph shows peak alignment as a function of Tone in the subsets where its influence can be compared in the same level of Stem length. The means for peak alignment differ by around 40 ms between Low<sup>Fall</sup> and Fall. The size of this difference is similar in Bor South and in Bor North. Moreover, in the Bor North dialect, where the Fall is also found on short vowels, the difference in peak alignment between the Low<sup>Fall</sup> and the Fall is similar in size between short and long stem vowels.

It is worthwhile to examine the peak alignment of the Low<sup>Fall</sup> in greater detail—this tone pattern was collected in six different length conditions, through the orthogonal crossing of Stem length and Prefix length; see 11 above. The result is a highly detailed perspective on the influence of time pressure on the alignment of tonal targets.

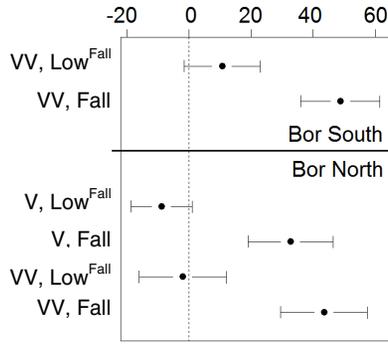


FIGURE 7. Means and standard deviations for peak alignment by Dialect, Tone, and Stem length (V and VV only).

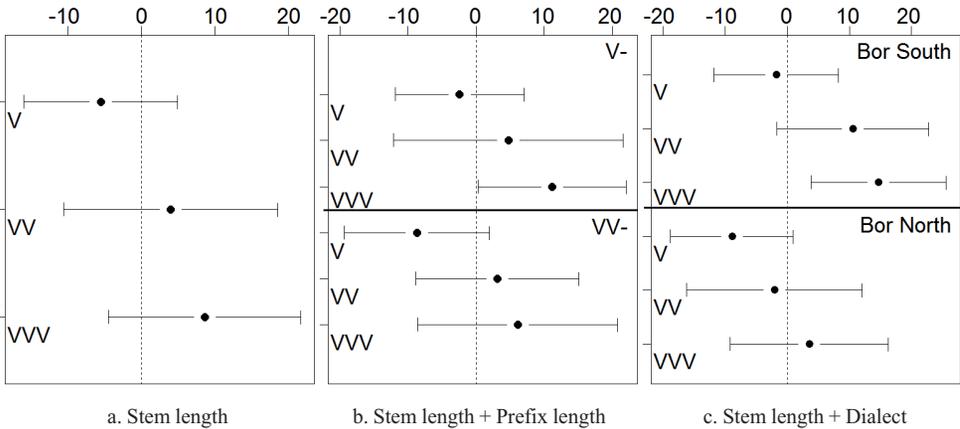


FIGURE 8. Means and standard deviations for peak alignment in the realization of Low<sup>Fall</sup>, as a function of various combinations of Stem length, Prefix length, and Dialect.

As seen from Figure 8a, the peak is aligned progressively later as Stem length increases from short (V) to long (VV), and further to overlong (VVV). In contrast, an increase in Prefix length has the opposite effect: from a short (V-) to a long (VV-) prefix vowel, peak alignment shifts earlier (Fig. 8b, upper panel vs. lower). In both cases, there is a shift in alignment toward the region where there is more segmental duration.

The third factor under consideration in Fig. 8 is Dialect. As seen from Fig. 8c, its effect is similar to that of Prefix length (Fig. 8b): the Low<sup>Fall</sup> reaches its high initial target earlier in Bor North than in Bor South. There is no obvious explanation for this difference, in particular given that a comparable difference can be observed in relation to the Fall (Fig. 7, Table 6).

A final factor to consider is Coda type. Three of the four sets have a sonorant coda (N); the fourth set has a plosive coda (P). Figure 9 shows the descriptive statistics. Peak alignment is earlier when the coda is a plosive than when it is a sonorant, across levels of Tone and Stem length. This difference is small in the case of the Low<sup>Fall</sup>, which is aligned 3.8 ms earlier when the coda consonant is a voiceless plosive rather than a sonorant, but is more substantial in the case of the Fall, where peak alignment is on average 7.5 ms earlier when the coda is a voiceless plosive. The difference is particularly great when the stem vowel is short and the syllable carries a Fall. In this structure, the one where time pressure is greatest, the peak is aligned 14.7 ms earlier when the coda is a voiceless plosive than when the coda is a sonorant.

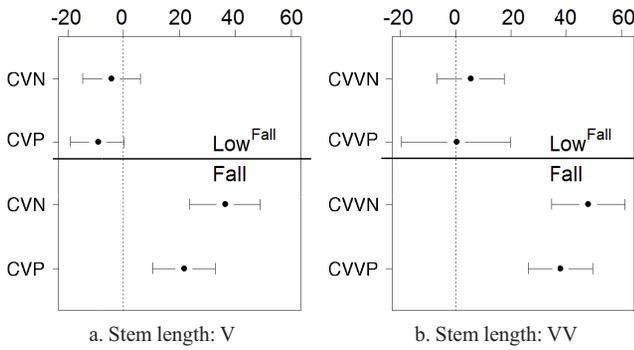


FIGURE 9. Means and standard deviations for peak alignment by Stem length (separate panels for short and long vowels), Tone, and Coda Type.

The direction of the difference in the peak alignment as a function Coda type can be explained in the same way as the differences observed for Stem length and Prefix length. When the coda is a plosive, the sonorous duration on which the tone pattern can be realized is smaller than when the coda is a sonorant. This reduction in sonorous duration shifts the peak earlier, away from the area where the segmental space is reduced. And given that a voiceless plosive gives rise to increased time pressure at the end of the target syllable, it is natural for the later-aligned Fall to be affected more than the earlier-aligned Low<sup>Fall</sup>.

In summary, Tone conditions a substantial difference in peak alignment, of about 40 ms on average, and the distributions for Low<sup>Fall</sup> and Fall do not overlap at the level of one standard deviation. All of the other factors (Stem length, Prefix length, Dialect, Coda type) give rise to smaller differences, of 5 to 10 ms per factor, and here there is substantial overlap between the distributions of factor levels.

Table 7 reports the results of the mixed-effects model with dependent peak alignment (expressed relative to the beginning of the stem vowel). Tone, Stem length, Dialect, and Prefix length all yield a significant effect, in that order of importance. Coda type does not condition a significant main effect, but its interaction with Tone is significant. This indicates that the significant influence of Coda type on peak alignment is limited to the Fall (cf. Fig. 9). Tone has the biggest effect on peak alignment, with a *t*-value of 31.3. The significant effects conditioned by other factors are much smaller in size, with *t*-values under 10.

FACTOR AND LEVELS	<i>t</i> -VALUE	<i>p</i>
Stem length: V vs. VV	7.5	< 0.001
V vs. VVV	9.5	< 0.001
VV vs. VVV	3.2	0.002
Prefix length: V- vs. VV-	-3.1	0.002
Coda type: sonorant vs. plosive	-1.3	0.181
Tone: Low <sup>Fall</sup> vs. Fall	31.3	< 0.001
Dialect: BS vs. BN	-3.7	< 0.001
Coda type * Tone:	-2.9	0.003

TABLE 7. The results of a linear mixed-effects model with dependent peak alignment.

PEAK HEIGHT. The influence of declination—that is, the overall downward slope of *f*<sub>0</sub> throughout the utterance—was controlled in the data set by keeping constant both the number of syllables in the sentences and also the tonal context. That is, the target—a di-

syllabic verb form—is in sentence-final position, preceded by a Low-toned monosyllabic noun. As a result, any difference in the height of the  $f_0$  peak at the beginning of the falling contour can be attributed to the factors under investigation. The traces in Fig. 6 do not reveal a salient difference between  $\text{Low}^{\text{Fall}}$  and Fall in the  $f_0$  value of the peak. This is confirmed by the mean values for peak height, which are only 2.1 Hz apart: 150.7 Hz for the  $\text{Low}^{\text{Fall}}$  vs. 148.6 Hz for the Fall.

The results of the inferential test with peak height as the dependent are presented in Table 8. Tone ( $\text{Low}^{\text{Fall}}$  vs. Fall) does not yield a significant effect. The only factor that does reach significance is Prefix length. The peak is slightly higher when the prefix vowel is short than when the prefix vowel is long. The mean values are 151.4 Hz and 148.7 Hz, respectively, that is, a difference of 2.7 Hz. This difference can be attributed to declination: when the vowel of the prefix syllable is long, the stem syllable appears slightly further along the declination slope, thereby yielding a lower peak value. The difference in peak height between  $\text{Low}^{\text{Fall}}$  and Fall, while not significant, can be explained in the same way:  $f_0$  has declined slightly further in the case of the later-aligned Fall (mean peak height: 148.6 Hz) than in the earlier-aligned  $\text{Low}^{\text{Fall}}$  (mean peak height 150.7 Hz).

FACTOR AND LEVELS		<i>t</i> -VALUE	<i>p</i>
Stem length:	V vs. VV	-0.20	0.820
	V vs. VVV	-0.10	0.999
	VV vs. VVV	0.20	0.853
Prefix length:	V- vs. VV-	-3.60	< 0.001
Coda type:	sonorant vs. plosive	1.62	0.106
Tone:	$\text{Low}^{\text{Fall}}$ vs. Fall	-0.50	0.635
Dialect:	BS vs. BN	-0.20	0.801

TABLE 8. Summary of the results of a linear mixed-effects model with dependent peak height.

**EXCURSION SIZE OF THE  $f_0$  FALL.** In the previous section, it became clear that the height of the  $f_0$  peak that defines the beginning of the falling contour varies little as a function of the factors under investigation. Given this, variability in the excursion size of the  $f_0$  fall is determined primarily by its end value. A comparison between the panels in Fig. 6 above suggests that the size of the  $f_0$  fall is reduced when the coda is a voiceless plosive and therefore does not carry  $f_0$ . In addition, Fig. 6a suggests that, specifically in relation to stem forms with a sonorant coda, a greater proportion of the  $f_0$  fall takes place in the coda consonant (i) when Stem length is reduced—from VVV to VV to V—and (ii) when the Tone involves later alignment—that is, from  $\text{Low}^{\text{Fall}}$  to Fall. These differences can be observed more clearly in Figure 10, which shows descriptive statistics for the excursion size of the fall. Figure 10a shows the results for syllables with a sonorant coda. Here there is little difference in the size of the  $f_0$  fall: the mean excursion sizes for the  $f_0$  fall as a function of Stem length and Tone are all close to 40 Hz. Figure 10b displays the size of the fall, again in syllables with a sonorant coda, but now only up to the end of the vowel. These statistics show that the size of the  $f_0$  fall over the vowel becomes smaller as Stem length decreases and as Tone involves later alignment. This is relevant, because the vowel is the most salient part of the rhyme in perceptual terms.

Figure 10c shows similar decreases in the size of the fall as a function of Stem length and Tone when the coda is an unvoiced plosive. When time pressure is maximal—that is, when the stem vowel is short, the tonal specification is the late-aligned Fall, and the coda is a voiceless plosive—the size of the  $f_0$  fall is reduced to 12.6 Hz. This amounts to 8.3% of the relevant peak value that defines the beginning of the fall (151.7 Hz).

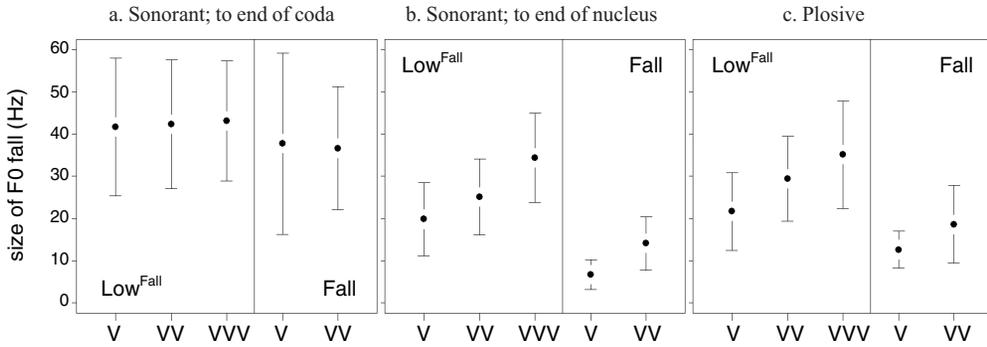


FIGURE 10. Means and standard deviations for the size of the f0 fall, as a function of Tone and Stem length, starting from the peak. Separate panels by Coda type (Sonorant, Plosive) and domain.

The results of the inferential test with the size of the f0 fall as the dependent are reported in Table 9. The most sizeable main effect is conditioned by Coda type. As seen from the negative *t*-value, the size of the f0 fall is smaller when the coda is a voiceless plosive than when it is a sonorant. There is also a main effect of Stem length, between short and both long and overlong vowels—that is, where time pressure is more acute—but not between long and overlong vowels. Prefix length also conditions a significant difference in fall size, although the effect size is small: the probability value is close to the significance threshold (0.05).<sup>8</sup>

FACTOR AND LEVELS		<i>t</i> -VALUE	<i>p</i>
Stem length:	V vs. VV	-2.60	0.011
	V vs. VVV	-3.40	< 0.001
	VV vs. VVV	1.30	0.190
Prefix length:	V- vs. VV-	-2.10	0.035
Coda type:	sonorant vs. plosive	-6.30	< 0.001
Tone:	Low <sup>Fall</sup> vs. Fall	-0.14	0.888
Dialect:	BS vs. BN	0.02	0.984
Tone * Coda type		-2.50	0.013
Stem length (V vs. VV) * Coda type		3.50	< 0.001
Stem length (V vs. VVV) * Coda type		5.10	< 0.001
Stem length (VV vs. VVV) * Coda type		-2.20	0.032

TABLE 9. Summary of the results of a linear mixed-effects model with excursion size of f0 fall over the voiced part of the rhyme as the dependent.

There are also two significant interactions, both involving Coda type. One involves Tone. This reflects the fact that Low<sup>Fall</sup> and Fall differ in the excursion size of the f0 fall more when the coda is a plosive (Fig. 10c) than when the coda is a sonorant (Fig. 10a). The second significant interaction involves Stem length. The size of the f0 fall varies as a function of Stem length particularly when the coda is a plosive (Fig. 10c).

In summary, the main effects and interactions show that the excursion size of the fall is affected by factors that have a bearing on the amount of sonorous duration in the stem syllable—in particular Stem length and Coda type—and by Tone, which affects the proportion of this sonorous domain over which the fall is realized.

<sup>8</sup> The effect of Prefix length on the excursion size of the fall may be due to the influence of Prefix length on the duration of the vowel of the stem syllable. That is, mean stem-vowel duration is slightly shorter after a long prefix vowel (115.6 ms) than after a short prefix vowel (117.5 ms). As seen from Table 3, this difference is not significant.

#### 4. DISCUSSION.

**4.1. EVALUATION OF THE COMPETING HYPOTHESES.** The null hypothesis (1) states that the alignment of falling contour tones is not phonologically distinctive. The alternative hypothesis (10) challenges this, specifically in relation to the surface phonology of Dinka.

Likelihood ratio tests were carried out to test these hypotheses directly, on the basis of the dependent measures examined in §3. In each test, the fully developed mixed-effects model is compared to a baseline model in which the factor Tone (Low<sup>Fall</sup> vs. Fall) and any interactions involving Tone are not included. The outcomes are reported in Table 10. The contrast between Low<sup>Fall</sup> and Fall significantly contributes to explaining the variability in tonal alignment, the size of the f0 fall, and duration, in that order of importance.

DEPENDENT	CHI-SQUARE	<i>p</i>
Alignment of peak relative to vowel onset	558.1	< 0.001
Excursion size of the fall	52.6	< 0.001
Vowel duration	10.8	< 0.001
Peak height	0.2	0.641

TABLE 10. Outcomes of likelihood ratio tests of the models reported in Tables 7, 9, 3, and 8, respectively, each relative to a baseline in which Tone (Low<sup>Fall</sup> vs. Fall) is not included.

Peak alignment is to be interpreted as the primary correlate of the contrast, on several grounds. As seen from the chi-square values in Table 10, this is the dependent that yields the biggest effect size as a function of the addition of Tone to the baseline model. In other words, the contrast between Low<sup>Fall</sup> and Fall affects peak alignment more than any other measure.

And whereas Tone influences peak alignment overall, that is, as a main effect, the effect of Tone on the size of the f0 fall is conditional on the composition of the target syllable. The addition of an interaction between Coda type and Tone significantly improves on a model in which no interaction is modeled, and, with this interaction included in the model, Tone is no longer significant as a main effect (Table 9). The third measure significantly affected by the addition of Tone to the baseline model is vowel duration. It is not uncommon in tone languages for tone patterns to have a secondary effect on duration (e.g. Ho 1976 on Mandarin). Finally, the addition of Tone to the model for peak height is not significant. This is in line with the descriptive statistics, which show a difference of only 2.1 Hz. This outcome dispels an interpretation whereby the contrast of Low<sup>Fall</sup> vs. Fall is analyzed in terms of tone height.

This outcome corroborates the alternative hypothesis, thereby confirming Andersen's (1987:20) descriptive analysis of contrastive alignment in falling contours in Dinka. This raises the question of phonological representation: how can contrastive alignment be best dealt with in the phonology? This issue is addressed in §4.3. First, however, the findings are considered in the light of phonetic limitations relating to tonal alignment. These limitations can usefully inform the question of phonological representation.

**4.2. PHONETIC LIMITATIONS ON ALIGNMENT CONTRASTS IN CONTOUR TONES.** Across the world's languages, the distribution of contour tones is often conditioned by factors determining the size of the sonorous rhyme duration (Zhang 2001, 2002)—factors such as vowel length, the presence and nature of a coda, stress, and phrase-final lengthening. Ultimately, the relevance of these factors follows from phonetic limitations on the realization of f0 contours and on the perception of the resulting pitch impressions. In this

section, the results of the production study in §3 are considered in relation to these limitations.

The duration of the interval over which a fall in  $f_0$  is realized varies greatly within the data set. Relevant statistics are presented in Table 11. The fourth column shows the mean durations from the  $f_0$  peak to the end of the sonorant domain, as a function of three factors that affect time pressure. The means range from 48.4 ms for the Fall on a syllable with a short vowel and a voiceless coda (CVP), up to a domain that is more than five times greater (274.4 ms) for the  $\text{Low}^{\text{Fall}}$  on a syllable with voiced coda and overlong vowel (CVVVN).

FACTORS AFFECTING TIME PRESSURE			DURATION OF F <sub>0</sub> FALL	SIZE OF F <sub>0</sub> FALL		SIZE OF F <sub>0</sub> FALL – GLISSANDO THRESHOLD
Tone	Coda type	Stem length		Hz	ST	
Fall	P	V	48.4	12.6	1.5	-1.8
Fall	P	VV	76.0	18.6	2.3	0.2
Fall	N	V	141.7	37.7	5	3.9
Fall	N	VV	157.7	36.6	4.8	3.8
$\text{Low}^{\text{Fall}}$	P	V	75.0	21.7	2.7	0.6
$\text{Low}^{\text{Fall}}$	P	VV	103.1	29.4	3.8	2.2
$\text{Low}^{\text{Fall}}$	P	VVV	178.2	35.1	4.6	3.7
$\text{Low}^{\text{Fall}}$	N	V	166.2	41.7	5.8	4.8
$\text{Low}^{\text{Fall}}$	N	VV	200.9	42.3	5.7	4.9
$\text{Low}^{\text{Fall}}$	N	VVV	274.1	43.1	5.8	5.2

TABLE 11. Mean values as a function of Tone, Coda type, and Stem length, for the duration of the  $f_0$  fall (ms) and size of the  $f_0$  fall (Hz and ST). The last column shows the difference between the mean size of the fall and the relevant glissando threshold, calculated using the formula in 12.

The duration of the domain over which the  $f_0$  fall is realized influences its excursion size. Relative to mean values of around 150 Hz for the  $f_0$  peak that defines the beginning of the  $f_0$  fall, the mean size of the fall ranges from 12.6 Hz (Fall on CVP) to an excursion size that is more than three times greater: 43.1 Hz ( $\text{Low}^{\text{Fall}}$  on CVVVN). This dependency of the size of the  $f_0$  fall on the duration of the  $f_0$  fall suggests that the speed of  $f_0$  change is limited (cf. Xu & Sun 2002). Xu and Sun point out that  $f_0$  changes in speech are shaped like S-curves; that is, there is a fast-moving central section that is bounded on either side by inert sections near the turning points. This can be seen from Fig. 6a: when the Fall is realized on a syllable with a short vowel, the inert section around the peak takes up most of the vowel's duration. When the vowel is long (VV), the speakers reach the fast-moving section well within the vowel. As for the bottom end of the falling  $f_0$  change, Figs. 6a and 10a show that the inert final section absorbs the effects of time pressure, thereby buffering the high-speed central section from truncation. This characteristic, along with final lengthening, can help to explain why the utterance-final position is conducive to the realization of contour tones (cf. Zhang 2001, 2002). The fast-moving central section is truncated only if the loss of sonorous duration is more substantial (Figs. 6b, 10c).

The reduced excursion size of the  $f_0$  fall under time pressure has implications for pitch perception, because  $f_0$  changes are perceived as pitch contours only if they exceed a threshold excursion size (Rossi 1971, Greenberg & Zee 1979, 't Hart et al. 1990).  $F_0$  changes below this 'glissando threshold' are perceived as pitch levels. On the basis of their own experiments and the results of several earlier studies, 't Hart and colleagues (1990:32) model the glissando threshold as in 12. As seen from this formula, the glissando threshold increases exponentially as the duration of the domain over which the  $f_0$

change is realized (*D*) is reduced. For example, while a change from 150 to 135 Hz (1.83 ST) may be perceived as a pitch contour over 100 ms, the same change may not be perceived as an *f*<sub>0</sub> change over 50 ms.<sup>9</sup>

$$(12) \text{ Glissando threshold (ST per second)} = 0.16 / D^2$$

This formula enables us to evaluate whether the *f*<sub>0</sub> falls in the production study are more likely to be perceived as levels or as contours. The fifth and sixth columns in Table 11 provide the size of the *f*<sub>0</sub> fall, and the seventh column has the size of the fall expressed relative to the glissando threshold, calculated on the basis of 12. In most conditions, the *f*<sub>0</sub> falls realizing *Low*<sup>Fall</sup> and *Fall* are two or more semitones above the threshold. But when the *Fall* is realized on CVP, the *f*<sub>0</sub> fall is 1.8 ST below the glissando threshold. In this context, the *Fall* would be perceived as a level pitch. In addition, the values for a *Fall* on CVVP and for the *Low*<sup>Fall</sup> on CVP are close to the threshold. A comparison of the values indicates that a plosive coda is more influential in bringing the size of the *f*<sub>0</sub> fall down toward the glissando threshold than a short vowel. This is due to the fact that a sonorant coda represents a greater contribution to sonorous space—*D* in 12—than an increase in vowel length. This helps to explain why a plosive conditions a leftward shift on peak alignment (see Fig. 9).

The evaluation of the size of *f*<sub>0</sub> falls in relation to the glissando threshold reveals that the range of pitch percepts of the *Fall* toneme probably constitutes a level pitch pattern. This level allophonic variant of the *Fall* is found in the Bor North dialect of Dinka, when the *Fall* is realized on a syllable with a short vowel and a voiceless coda. Ohala (1989) has argued that diachronic change in sound patterns is ‘drawn from a pool of synchronic variation’. One of the main mechanisms in this context is hypocorrection, whereby the listener fails to correct for a contextual perturbation. In this way, the level tone percept of the *Fall* could be reinterpreted as reflecting a level tone pattern in a phonological sense. This scenario of hypocorrection is precisely what appears to have happened in Bor South: here the *Mid* toneme is found synchronically on a short vowel in contexts where other Dinka dialects (Bor North, Luanyjang, Agar) have the *Fall* toneme. Figure 11 contrasts the *Fall* of Bor North with the *Mid* of Bor South, as a function of Coda type. The data are the passive forms of verbs with a short vowel (see Table 1). When the *Fall* of Bor North is maximally truncated (solid black line), the acoustic difference with the *Mid* of Bor South (solid gray line) is small. Perceptually, the difference would be smaller still, because the level-tone pitch percept of a frequency change is skewed toward the end value (Nábělek et al. 1970, Rossi 1971). In conclusion, the truncated realization of the *Fall* in Bor North is similar to its cognate in Bor South.

A third phonetic consideration is the sensitivity to the difference in timing between *f*<sub>0</sub> patterns. Most perception tests relating to this issue report that it takes a difference of at least 50 ms in timing relative to the segmentals for two otherwise identical sequences of tone targets to be reliably distinguished (House 2004 and further references there). However, Bruce (1977) reports a timing sensitivity below 30 ms, albeit on the basis of stimuli that may not be ecologically valid.<sup>10</sup> In the Dinka data under investigation here,

<sup>9</sup> If *D* is 100 ms (0.1 second), the glissando threshold (*G*) is 16 ST per second, or 1.6 ST over 100 ms. If *D* is 50 ms (0.05 second), then *G* is 64 ST per second, or 3.2 ST over 50 ms. Hence a change of 1.83 ST is above the glissando threshold for a duration of 100 ms, but below it for a duration of 50 ms.

<sup>10</sup> Bruce’s stimuli involve falls from 160 to 100 Hz (8 ST) over 40/60/80 ms. In contrast, the fastest speaker (out of thirty-three subjects) in the study by Xu and Sun (2002:1407) takes at least 97 ms to produce an *f*<sub>0</sub> change of this size.

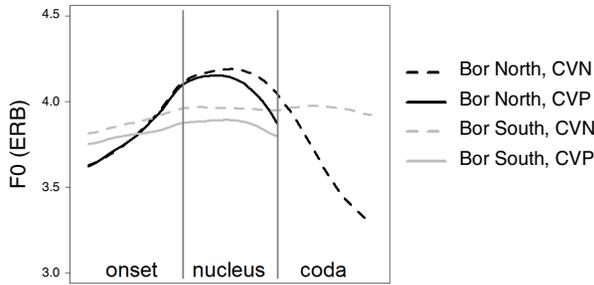


FIGURE 11. Averaged f0 traces on a normalized time axis of the Fall of Bor North and the Mid of Bor South, on syllables with a short stem vowel. Separate traces by Coda type and Dialect.

the differences in peak alignment between Low<sup>Fall</sup> and Fall all lie between 30 and 50 ms (Table 12). The difference is smallest when time pressure is maximal: in the Bor North dialect in CVP forms. As noted above, however, this contrast may be one of a pitch contour vs. a level pattern, in perceptual terms.

DIALECT	CONTEXT	ALIGNMENT DIFFERENCE Low <sup>Fall</sup> vs. Fall
Bor North	CVVC	45.6 ms
	CVN	41.4 ms
	CVP	30.9 ms
Bor South	CVVC	38.2 ms

TABLE 12. Differences between mean peak alignment as a function of Tone (Low<sup>Fall</sup> vs. Fall), in various conditions.

The small-scale perception experiment reported in §2.2 suggests that Dinka speakers can use these patterns of alignment to distinguish lexical and grammatical meanings. It is not implausible that the use of contrastive alignment in Dinka makes its speakers more sensitive to differences in tonal alignment, as compared to speakers of languages where this parameter is not exploited to the same extent in the phonology. Such a state of affairs would be in line with the fact that linguistic experience of word-level tonal specification in a first language influences pitch perception (e.g. Gandour 1983, Pfordresher & Brown 2009).

In summary, I have considered the results of contrastive alignment in Dinka against the background of experimental findings on the phonetics of tone. It is clear that the contrast in tonal alignment between falling contours in Dinka presents a phonetic challenge. First, the observed differences in alignment (30 to 50 ms) are low relative to the range reported in perception experiments. Second, the evaluation of the f0 falls relative to the glissando threshold reveals the difficulty in realizing the late-aligned fall as a contour in perceptual terms. The perceptual challenge is not only to distinguish the two falling contours from one another, but also to distinguish the late-aligned fall from any level tone categories. The cross-dialect comparison confirms that the late-aligned fall can be reinterpreted as a level tone pattern.

These findings are relevant to the upcoming discussion of phonological representation: they suggest that it would be highly unlikely for a language to distinguish more than two falling contour tone categories primarily through the alignment of the same tone targets.

**4.3. THE PHONOLOGICAL REPRESENTATION OF CONTRASTIVE ALIGNMENT.** How should the contrast between  $\text{Low}^{\text{Fall}}$  and Fall be dealt with in terms of phonological representation? A first approach to consider is not to phonologically represent the variation in the realization of the Low toneme at all, that is, to leave it to phonetic implementation. This is the approach taken in Remijsen & Ladd 2008 in relation to Luanyjang, another dialect of Dinka. Such an analysis is plausible if the initial high target of the  $\text{Low}^{\text{Fall}}$  is predictable from the preceding tonal context. The Luanyjang tone category that corresponds to the Mid toneme of Bor Dinka is realized as a rising contour in utterance-medial contexts. It can be interpreted as a Rise (underlying Low-High), with a High end target that is realized in the following syllable, triggering the  $\text{Low}^{\text{Fall}}$  variant of a following Low.

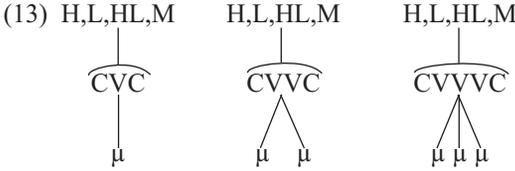
In relation to the Bor data under investigation here, however, such an account is problematic for several reasons. First, there is no independent evidence that the Mid involves a high target in its implementation. That is, when the Mid toneme is not followed by the  $\text{Low}^{\text{Fall}}$ , it is realized as a level tone pattern, as in Figs. 2 and 11. Second, if the initial high target of the  $\text{Low}^{\text{Fall}}$  were due to a pattern associated with the preceding syllable, then we would expect the alignment of this high target to be determined more by vowel length in this preceding syllable. The descriptive statistics in Tables 4 and 5 and the inferential statistics in Table 7 indicate that the opposite is the case: the high target of the  $\text{Low}^{\text{Fall}}$  is aligned most consistently with the beginning of the vowel of the stem syllable. Finally, phonetic implementation cannot account for all of the contexts in which  $\text{Low}^{\text{Fall}}$  is found. Its distribution includes the isolation form (Fig. 2a), where there is no earlier syllable with which the high target of the  $\text{Low}^{\text{Fall}}$  could be associated. The divergence in the initial  $f_0$  value between words specified for Mid vs. Low toneme suggests that there is no intonational edge tone involved here.

In conclusion, an account based on phonetic implementation is not satisfactory. Given this, the alignment contrast between  $\text{Low}^{\text{Fall}}$  and Fall needs to be represented somehow in the surface phonology. The autosegmental approach to tone and intonation offers two general strategies to achieve this. One is to associate tone units to more specific elements on the metrical tier. The other is to develop the specification of the tonal units. These two analytic options are foreshadowed by Pierrehumbert and Beckman (1988:159), who write: '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'. I consider these approaches in turn in the following subsections.

**REPRESENTING CONTRASTS ON THE PROSODIC NODES.** The metrical unit below the syllable is the syllable-internal weight unit or mora (Hyman 1985, Hayes 1989). In the study of tone, the mora is commonly invoked as the tone-bearing unit (TBU) when the distribution of tone patterns is richer in heavy syllables than in light ones (Odden 1995, Yip 2002, Gussenhoven 2004). For example, Myers (2003) reports that, in Kinyarwanda, Low contrasts with High on syllables with a short vowel, but on syllables with a long vowel Low contrasts with two tone patterns. One is falling; the other is high or rising. Myers represents these facts by postulating that short and long vowels project one and two moras, respectively, and that the mora is the TBU. It follows that a High target can associate with the first or with the second mora of a long vowel, yielding HL and LH, respectively, and that only one point of association is available on a syllable with a short vowel.

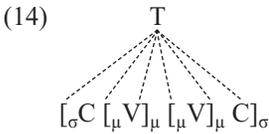
Can the Dinka contrast between  $\text{Low}^{\text{Fall}}$  and Fall be represented through association with moras? The three-level vowel length indicates that syllable-internal weight struc-

ture matters in Dinka, and this contrast can be represented by postulating representations involving one, two, and three moras for short, long, and overlong vowels respectively. In many dialects, however, including Bor North, the distribution of tone patterns does not interact with syllable weight: level and contour tones alike are found on syllables with a short, long, or overlong vowel, as in 13. This implies that the syllable as a whole is the TBU (Odden 1995, Yip 2002, Gussenhoven 2004).

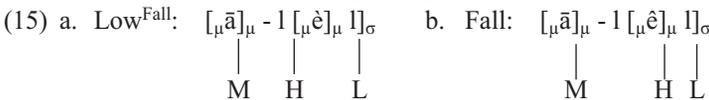


In his argument for an alternative criterion to determine metrical structure within the syllable, Duanmu (1994) proposes that the association of two tone targets with the same tone-bearing unit is ruled out, so that contour tones necessarily motivate two moraic TBUs. This analysis is problematic for Dinka, given that there are minimal contrasts such as /lêl/ vs. /lêel/ vs. /lêeel/: if we invoke two moras in the case of /lêl/ so as to represent the contour tone, then by extension we end up with four moras for /lêeel/. This is typologically undesirable (e.g. Odden 2011). Moreover, even if we had two TBUs in stem syllables with a short vowel, this still does not enable us to differentiate between two patterns that both involve a High target followed by a Low one. In summary, if contrastive alignment is to be represented by making metrical specification more specific, merely postulating the mora as a TBU does not present a solution.

Two studies—Prieto et al. 2005 and Morén & Zsiga 2006—have independently invoked a more detailed representation, whereby tone targets are associated with the left and right ‘edges’ of moras. This proposal hinges on an interpretation of the mora as a concrete phonetic constituent. For example, Morén and Zsiga (2006:126) postulate for Thai that, in a syllable whose rhyme includes a long vowel, a sonorant coda, or both, the right edge of the first mora lies halfway through the phonetic duration of the rhyme. The contrast between the Fall and High tonemes of Thai is represented through an H target associated with the first vs. the second mora, respectively, in each case in relation to the mora’s right edge. Prieto and colleagues make the case for alignment with mora edges more generally. For a tone target within a syllable with a long (bimoraic) vowel, their model provides the possible docking sites in 14: a tone can be linked with either edge of the syllable or of the moras within it (Prieto et al. 2005:373ff.).



The innovation of Prieto et al. 2005 and Morén & Zsiga 2006 expands the set of possible patterns of association, so that one mora offers two points of association rather than just one. Applied to Dinka, this enables us to represent the two falling tone patterns as in 15. The high initial targets of Low<sup>Fall</sup> and the Fall are represented through association with the mora’s left and right edges, respectively.



This approach hinges on an interpretation of the mora as a concrete timing unit. As such it can be tested on the basis of the results of the time-pressure study. In the case of the

Low<sup>Fall</sup>, the representation in 15a is in line with the phonetic data: the f<sub>0</sub> peak at the beginning of the falling contour is aligned closest to the beginning of the stem vowel. In the case of the Fall, however, it runs into problems. If the vowel is short (monomoraic), we would expect the f<sub>0</sub> peak to be reached at the end of the vowel, that is, at 100% of the mora's domain. Instead, the peak is reached about 44% into this domain: at 32.6 ms into a vowel of 74.1 ms. Contrary to the representation in 15b, the peak of the Fall is aligned closer to the left edge of the stem vowel than to its right edge. In summary, the phonetic evidence does not support an interpretation whereby the mora is used as a physical timing unit with whose edges the tone targets are lined up in a predictable way.

Detailed acoustic studies on tonal alignment in other languages are equally challenging for an interpretation whereby tones associate with the edges of moras. First, the alignment of tone targets often varies greatly across contexts. For example, the LH pattern of Kinyarwanda reaches its peak 25% earlier in the syllable in phrase-final position than phrase-medially. Based on these and other instances of variability in Kinyarwanda, Myers (2003:94) concludes that 'the mora is not [to be] treated as a unit of time, or as a specific stretch of the soundwave'. Second, the tone targets of a contour may be aligned well outside of the relevant domain. In a time-pressure study on Mandarin Chinese, Xu (2001) reports that the High and Rise tonemes both vary in peak alignment across the right edge of the syllable with which they are underlyingly linked. This is unexpected, at least if the turning points are interpreted as an accurate reflection of the phonological tone targets.<sup>11</sup>

A more fundamental problem for moraic alignment is that it is not sufficiently restrictive. As seen in 14, Prieto and colleagues (2005) allow for three separate representations of a bitonal contour in a bimoraic syllable: the initial target can be linked with the left edge of the first mora, the right edge of the first mora, or the right edge of the second mora. While the current study provides evidence for a contrast between early and late alignment (§4.1), it equally provides evidence of limitations to the production and perception of such tone patterns (§4.2). For example, the f<sub>0</sub> fall involved in the realization of the Fall on a bimoraic CVVP syllable is only 0.2 ST above the glissando threshold. If a third falling pattern were to be fitted into the same domain, it would be below the glissando threshold consistently. These limitations suggest that an inventory involving three contours, of the same shape but diverging in alignment, is beyond the range of phenomena a theory of tonal alignment needs to be able to accommodate.

REPRESENTING ALIGNMENT ON THE TONAL TIER. Apart from the proposal to associate tone targets with mora edges, autosegmental theory does not offer a way to represent a contrast involving contours of the same shape (see Pierrehumbert & Beckman 1988:159, cited above). For example, contrasts distinguished through the starredness convention (e.g. H\*+L vs. H+L\*) involve a sequence of tone targets that are realized over consecutive syllables, rather than two targets within the same syllable domain (Beckman & Pierrehumbert 1986:257). Other proposals relating to the phonological representation of tone targets cannot serve this purpose either. Ladd (1983) proposes a feature [ $\pm$ delayed peak], but this distinguishes high or falling accents from rising accents (Ladd 1983:730), rather than contours of the same shape. Akinlabi and Liberman (2001) argue for a tonal complex with two slots, each of which can accommodate a tone target. All of these proposals are in line with the null hypothesis, whereby languages can have only one falling contour within the syllable.

<sup>11</sup> Xu (2001) argues for an alternative theory, in which f<sub>0</sub> slopes may also reflect tone targets.



If an  $f_0$  change that is identical in direction and excursion size sets in during the onset consonant or at the beginning of the vowel, as in the gray trace in Fig. 12, the resulting percept would be categorically different: it would be perceived in terms of level targets, with the end target likely to predominate.<sup>12</sup>

According to this interpretation, the perceptual space for tonal alignment is not homogeneous, with the distance in timing between two tonal configurations of the same shape within a syllable being the only factor determining whether the difference between them is perceptually salient. Instead, it is argued that a contrast in alignment within a syllable will be discriminated more easily if the constituent patterns are separated by the threshold. There is independent evidence that the perception of  $f_0$  changes is indeed different across the syllable. Verhoeven (1994) carried out discrimination tests involving continua of equidistant alignment patterns. In one type of continuum the subjects compared stimuli in which the  $f_0$  change took place in the onset or at the start of the vowel (early continuum); in the other type of continuum the  $f_0$  change set in further along in the syllable, in the course of the vowel (late continuum). Comparable continua were created for falling and for rising  $f_0$  changes. For rises and falls alike, discrimination was significantly more accurate for pairs of stimuli from the late continua.

House's finding that the alignment of  $f_0$  changes triggers a categorical shift in perception is supported by the intuitions of specialists on tone and intonation. For example, Hyman (2010:203) discusses a rising  $f_0$  trace on a disyllabic word as follows: '[T]he second syllable seems to have a continuous rise in it, suggesting maybe a L-LH transcription. Listening to it, however, it was clear that it was perceptually [L-H]. To have been L-LH, the transition from L to H would have had to take place later in the [second] syllable'. Hyman continues to make the same argument for a falling  $f_0$  trace, again contrasting the acoustic shape as an  $f_0$  change with a level pitch impression. Hyman's account is in line with House's hypothesis that it takes late alignment for an acoustic contour to be saliently perceived as a pitch movement.

The findings on contrasting falling contours in Dinka support the case for a quantal threshold in tonal alignment, because the category boundary is in the region that House (1990) hypothesized to be critical in determining the pitch percept. In this way, the feature [ $\pm$ late-aligned] is adequate in an explanatory sense.

**5. CONCLUSION.** The results of this investigation into falling tone contours in Dinka indicate that a language can contrast two falling contours, diverging primarily in the alignment of the  $f_0$  change within a syllable. This finding contradicts the widely held assumption that tonal alignment is not contrastive within the syllable (Odden 1995:450, Silverman 1997:479–80, Yip 2002:29).

The time-pressure study also indicates that this alignment contrast between falling contours is difficult to maintain in the face of phonetic limitations. First, the  $f_0$  change of the late-aligned pattern (Fall) is below the glissando threshold when the amount of sonorous space is maximally restricted, and in some other contexts the  $f_0$  change is only just above this threshold. This means that the Fall toneme encompasses a level pitch percept within its allophonic range. The interpretation that the Fall would be perceived as a level tone pattern is borne out in Bor South, where this toneme has been reinterpreted as a Mid toneme on short vowels, that is, when time pressure is greater (cf. Ohala 1989). Second, the acoustic difference in alignment between the two falling contours is at the bottom end of the range of values reported for sensitivity to timing differences in

<sup>12</sup> In the perception of the pitch level of frequency changes, the influence of the end frequency predominates (Nábělek et al. 1970:538, Rossi 1971).

perception studies (House 2004). The relevance of time pressure in Dinka is in line with crosslinguistic observations about the dependency of contour tone realization on the availability of sonorous space (Zhang 2001, 2002).

The evidence of phonetic limitations is relevant to the question of phonological representation: given the difficulty of producing and perceiving binary contrast in tonal alignment between contour tones, it is unlikely that a system of contrasts involving three or more levels of tonal alignment could be maintained. Based on this consideration, I have argued for a representation of tonal alignment through a feature ([±late-aligned]), that is, a representation that inherently imposes a maximum of two levels. This consideration is also part of the argument against a representation with reference to mora edges, which allows for three or more levels of contrast (e.g. Prieto et al. 2005). Finally, the location of the category boundary in Dinka, within the first 30 ms of the vowel, is in line with the model of House 1990. This suggests that the alignment feature invoked for Dinka may exploit a quantal threshold (Stevens 1989, Stevens & Keyser 2010).

Aside from contrastive tonal alignment, the current study also presents evidence of another suprasegmental phenomenon that is typologically unusual. Three-level vowel length in Dinka was first reported in Andersen 1987, and has been corroborated on the basis of phonetic evidence (Remijsen & Gilley 2008). The current study strengthens the case for three-level vowel length in the surface phonology of Dinka in two ways. First, it widens the range of dialects for which three-level vowel length has been reported. There is now evidence of three-level vowel length in three of the four dialect clusters, namely Agar (Andersen 1987, 1993), Rek (Remijsen & Gilley 2008), and Bor (this study). Second, the results provide additional phonetic evidence, based on sets for three-level vowel length in which tone is kept constant.

#### REFERENCES

- AKINLABI, AKIN, and MARK LIBERMAN. 2001. Tonal complexes and tonal alignment. *North East Linguistic Society (NELS)* 31.1–20.
- ANDERSEN, TORBEN. 1987. The phonemic system of Agar Dinka. *Journal of African Languages and Linguistics* 9.1–27.
- ANDERSEN, TORBEN. 1993. Vowel quality alternation in Dinka verb inflection. *Phonology* 10.1–42.
- BATES, DOUGLAS. 2005. Fitting linear mixed models in R—Using the lme4 package. *R News* 5.1.27–30.
- BECKMAN, MARY E., and JANET B. PIERREHUMBERT. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3.255–309.
- BLUMSTEIN, SHEILA E., and KENNETH N. STEVENS. 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America* 67.648–62.
- BOERSMA, PAUL, and DAVID WEENINK. 2010. Praat: Doing phonetics by computer. Version 5.1.43. Online: <http://www.praat.org/>.
- BRUCE, GÖSTA. 1977. *Swedish word accents in sentence perspective*. (Travaux de l'institut de linguistique de Lund 12.) Lund: CWK Gleerup.
- CASPERS, JOHANNEKE, and VINCENT J. VAN HEUVEN. 1993. Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. *Phonetica* 50.161–71.
- D'IMPERIO, MARIAPAOLA, and DAVID HOUSE. 1997. Perception of questions and statements in Neapolitan Italian. *Proceedings of Eurospeech 1997*, Rhodes, Greece, 251–54.
- DUANMU, SAN. 1994. Against contour tone units. *Linguistic Inquiry* 25.4.555–608.
- EDMONDSON, JEROLD A., and JOHN H. ESLING. 2006. The valves of the throat and their functioning in tone, vocal register and stress. *Phonology* 23.157–93.
- FLACK, KATHRYN. 2007. Templatic morphology and indexed markedness constraints. *Linguistic Inquiry* 38.4.749–58.

- FROTA, SÓNIA. 2002. Tonal association and target alignment in European Portuguese nuclear falls. *Laboratory phonology 7*, ed. by Carlos Gussenhoven and Natasha Warner, 387–418. Berlin: Mouton de Gruyter.
- GANDOUR, JACK. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics* 11.149–75.
- GORDON, RAYMOND G. (ed.) 2005. *Ethnologue: Languages of the world*. 15th edn. Dallas: SIL International.
- GREENBERG, STEVEN, and ERIC ZEE. 1979. On the perception of contour tones. *UCLA Working Papers in Phonetics* 45.150–64.
- GUSSENHOVEN, CARLOS. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- HAYES, BRUCE. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20.2.253–306.
- HO, AICHEN T. 1976. The acoustic variation of Mandarin tones. *Phonetica* 33.353–67.
- HOUSE, DAVID. 1990. *Tonal perception in speech*. (Travaux de l'institut de linguistique de Lund 24.) Lund: Lund University Press.
- HOUSE, DAVID. 1996. Differential perception of tonal contours through the syllable. *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP 96)*, Philadelphia, 2048–51.
- HOUSE, DAVID. 2004. Pitch and alignment in the perception of tone and intonation. *From traditional phonology to modern speech processing*, ed. by Gunnar Fant, Hiroya Fujisaki, Jianfen Cao, and Yi Xu, 189–204. Beijing: Foreign Language Teaching and Research Press.
- HERMES, DIK J., and JOOST C. VAN GESTEL. 1991. The frequency scale of speech intonation. *Journal of the Acoustical Society of America* 90.97–102.
- HYMAN, LARRY M. 1985. *A theory of phonological weight*. Dordrecht: Foris.
- HYMAN, LARRY M. 2010. How to study a tone language, with exemplification from Oku (Grassfields Bantu, Cameroon). *UC Berkeley Phonology Lab Annual Report 2010*, 179–209.
- KIPARSKY, PAUL. 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2. 85–138.
- LADD, D. ROBERT. 1983. Phonological features of intonational peaks. *Language* 59.4.721–59.
- MORÉN, BRUCE, and ELIZABETH ZSIGA. 2006. The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory* 24.113–78.
- MYERS, SCOTT. 2003. F0 timing in Kinyarwanda. *Phonetica* 60.71–97.
- NÁBĚLEK, IGOR V.; ANNA K. NÁBĚLEK; and IRA J. HIRSH. 1970. Pitch of tone bursts of changing frequency. *Journal of the Acoustical Society of America* 48.536–53.
- NOLAN, FRANCIS. 2003. Intonational equivalence: An experimental evaluation of pitch scales. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, 771–74.
- ODDEN, DAVID. 1995. Tone: African languages. *The handbook of phonological theory*, ed. by John A. Goldsmith, 444–75. Oxford: Blackwell.
- ODDEN, DAVID. 2011. The representation of vowel length. *The Blackwell companion to phonology*, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 465–90. Oxford: Wiley-Blackwell.
- OHALA, JOHN J. 1989. Sound change is drawn from a pool of synchronic variation. *Language change—Contributions to the study of its causes* (Trends in linguistics 43), ed. by Leiv E. Breivik and Ernst H. Jahr, 173–98. Berlin: Mouton de Gruyter.
- PFORDRESHER, PETER Q., and STEVEN BROWN. 2009. Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception & Psychophysics* 71.1385–98.
- PIERREHUMBERT, JANET B., and MARY E. BECKMAN. 1988. *Japanese tone structure*. (Linguistic inquiry monograph 15.) Cambridge, MA: MIT Press.
- PRIETO, PILAR; MARIAPAOLA D'IMPERIO; and BARBARA GILI FIVELA. 2005. Pitch accent alignment in Romance: Primary and secondary associations with metrical structure. *Language and Speech* 48.359–96.
- R DEVELOPMENT CORE TEAM. 2010. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: <http://www.R-project.org/>.

- REMIJSEN, BERT, and LEOMA GILLEY. 2008. Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). *Journal of Phonetics* 36.318–44.
- REMIJSEN, BERT, and D. ROBERT LADD. 2008. The tone system of the Luanyjang dialect of Dinka. *Journal of African Languages and Linguistics* 29.2.149–89.
- REMIJSEN, BERT, and CAGUOR A. MANYANG. 2009. Luanyjang Dinka. *Journal of the International Phonetic Association* 39.1.113–24.
- ROETTGER, LARRY, and LISA ROETTGER. 1989. A Dinka dialect study. *Occasional Papers in the Study of Sudanese Languages* 6.1–64.
- ROSSI, MARIO. 1971. Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 23.1–33.
- SCHEPMAN, ASTRID; ROBIN LICKLEY; and D. ROBERT LADD. 2006. Effects of vowel length and ‘right context’ on the alignment of Dutch nuclear accents. *Journal of Phonetics* 34.1–28.
- SILVERMAN, DANIEL. 1997. Tone sandhi in Comaltepec Chinantec. *Language* 73.473–92.
- STEVENS, KENNETH N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17.3–45.
- STEVENS, KENNETH N., and SAMUEL J. KEYSER. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics* 38.10–19.
- ’T HART, JOHAN; RENÉ COLLIER; and ANTONIE COHEN. 1990. *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- VERHOEVEN, JO. 1994. The discrimination of pitch movement alignment in Dutch. *Journal of Phonetics* 22.65–85.
- XU, YI. 1998. Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica* 55.179–203.
- XU, YI. 1999. Effects of tone and focus on the formation and alignment of  $f_0$  contours. *Journal of Phonetics* 27.55–105.
- XU, YI. 2001. Fundamental frequency peak delay in Mandarin. *Phonetica* 58.26–52.
- XU, YI, and XUEJING SUN. 2002. Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 111.1399–413.
- YIP, MOIRA. 2002. *Tone*. Cambridge: Cambridge University Press.
- YU, KRISTINE M. 2008. The prosody of second position clitics and focus in Zagreb Croatian. Los Angeles: University of California, Los Angeles master’s thesis. Online: [http://www.linguistics.ucla.edu/general/matheses/Yu\\_UCLA\\_MA\\_2008.pdf](http://www.linguistics.ucla.edu/general/matheses/Yu_UCLA_MA_2008.pdf).
- ZHANG, JIE. 2001. *The effects of duration and sonority on contour tone distribution—Typological survey and formal analysis*. Los Angeles: University of California, Los Angeles dissertation.
- ZHANG, JIE. 2002. *The effects of duration and sonority on contour tone distribution: Typological survey and formal analysis*. New York: Routledge.

[b.remijsen@ed.ac.uk]

[Received 26 September 2011;  
revision invited 14 June 2012;  
revision received 20 December 2012;  
accepted 9 January 2013]