

# The intonation of yes–no questions in Luganda

Scott Myers 

The University of Texas at Austin

[s.myers@austin.utexas.edu](mailto:s.myers@austin.utexas.edu)

The intonation of yes–no questions in Luganda (Bantu, Uganda) has only been sketched in passing. Hyman states that Luganda yes–no questions are marked by a ‘super-high tone’ immediately following the last lexical high tone in the sentence, but there is little agreement in the literature about the intonation of yes–no questions if there is no lexical high tone in the sentence. To clarify the differences between statements and yes–no questions in Luganda, an acoustic production study was conducted. Nineteen speakers read aloud sentences differing in the location of the last lexical high tone relative to the end of the sentence. Each sentence was produced as a statement and as a question. Analysis of  $f_0$  measurements supported Hyman’s description of sentences with a lexical high tone, since the questions had an  $f_0$  peak that was higher and later than in the corresponding statements. For sentences without a lexical high tone, yes–no questions were found to begin with an interval in which  $f_0$  is higher than in corresponding statements, and end with a final  $f_0$  value lower than in statements. It is proposed that the yes–no question marker is a phrase accent ( $H^-$ ). Like the high phrase accent posited by Pierrehumbert for English, this intonational tone is associated after the last tone in the phrase, but in Luganda that last tone is lexical, rather than being an intonational focus marker as in English. This  $H^-$  accent is subject to upstep in the position after a high tone.

## 1 Introduction

In tone languages, differences in the  $f_0$  trajectory distinguish one lexical item from another. It is therefore a challenge in tone languages to accommodate intonational tones, since these are also differences in  $f_0$  trajectory, distinguishing one sentence type from another. Hyman & Monaka (2011) note that in some tone languages, such as Coreguaje (Tukanoan, Colombia), the intonational tones marking statements and questions override and replace lexical tones. In other tone languages, intonational tones are just added to the sequence of lexical tones. For example, intonational boundary tones are placed in the last syllable of a phrase following all lexical tones in Swedish (Bruce 1977), Japanese (Pierrehumbert & Beckman 1988), Kinande (Hyman 1990), Chichewa (Myers 1996), and Akan (Genzel & Kügler 2020).

The yes–no question construction in Luganda, a Bantu tone language spoken in Uganda, provides an interesting case of the interaction of lexical tone with intonational tones. The lexical tone patterns of the language have been thoroughly documented in work by Tucker (1962), Cole (1967), Stevick (1969a), Hyman (1982), and Hyman & Katamba (1993, 2010), but this extensive literature contains only brief mentions of intonation in Luganda.

Hyman (1982: 28) states that in a sentence in which there is a lexical high tone, the yes–no question is marked by ‘a super-high interrogative tone’ immediately following that high tone, as in (1b). This interrogative tone is absent in the corresponding statement in (1a).

The standard orthography does not mark tone, so here and henceforth the orthographic representations of the examples are augmented with acute accents indicating lexical high tones.<sup>1</sup>

- |                         |                           |
|-------------------------|---------------------------|
| (1) a. [òmùpákàsi]      | b. [òmùpákási]            |
| o-mu-pákas-i            | o-mu-pákas-i              |
| IV-1-work.for.hire-NMLZ | IV-1-work.for.hire-NMLZ.Q |
| ‘A porter.’             | ‘A porter?’               |

If there is no lexical high tone in the sentence, according to Hyman (1982), there is a super-high question-marking tone on the second syllable, as in (2b).

- |                     |                       |
|---------------------|-----------------------|
| (2) a. [òmùlimì]    | b. [òmúlimì]          |
| o-mu-lim-i          | o-mu-lim-i            |
| IV-1-cultivate-NMLZ | IV-1-cultivate-NMLZ.Q |
| ‘A farmer.’         | ‘A farmer?’           |

However, in later work, Hyman (1990: 122) and Hyman & Katamba (2011: 71–72) provide a different description of questions with no lexical high tones, stating that such a question is low-toned throughout, as in (3).

- |                           |
|---------------------------|
| (3) [àbàgùlilà]           |
| a-ba-gul-ilil-a           |
| 3SG-3PL.OBJ-buy-APPL-FV.Q |
| ‘Is he bribing them?’     |

The brief description of yes–no questions given by Stevick (1969a: 27) states the case in which there is a lexical high tone in a way similar to Hyman (1982), except that he does not say that the intonational tone is super-high. In the case in which the final word in the sentence has no lexical high tone, Stevick says that ‘the final syllable is extremely low in pitch’.

These descriptions do not agree on some rather important factual matters: whether the intonational tone marking yes–no questions is higher in pitch than a lexical high tone in the same position, and whether there is such an intonational high tone in yes–no questions that have no lexical high tones. There are also some cases that they do not cover, such as sentences with more than one lexical high tone, or ones in which the last lexical high tone in the sentence is earlier than the final word.

However, these descriptions are in agreement that the question-marking high tone in Luganda is not limited to the final syllable of an intonational phrase, which insures that it is not a boundary tone (H%), in the sense of Pierrehumbert (1980) or Beckman & Hirschberg (1994). It is instead parallel in distribution to the phrase accents (H<sup>-</sup> and L<sup>-</sup>) of Pierrehumbert (1980), intonational tones which occur immediately following the final pitch accent in the phrase in English. The last lexical high tone in Luganda is not a pitch accent in the sense of Pierrehumbert (1980), since the syllable it is associated with is not stressed, but it is parallel

<sup>1</sup> Glossing abbreviations are as in the Leipzig Glossing Rules (<http://www.eva.mpg.de/lingua/resources/glossing-rules.php>) except for: 1–23 = noun class, FV = final vowel (terminal verb suffix), IV = initial vowel (i.e. augment).

to the nuclear accent in English in being the last tone before the end of the phrase. This class of intonational tone is modelled after the ‘sentence accent’ that Bruce (1977) posited for Swedish. Grice, Ladd & Arvaniti (2000) survey a number of intonational patterns in European languages in which an intonational tone occurs in this zone after the last pitch accent. It will be proposed in this paper that yes–no questions in Luganda are marked by a H<sup>-</sup> phrase accent.

The positioning of the question-marking high tone in Luganda yes–no questions is unusual from a comparative perspective. A high boundary tone on the phrase-final syllable distinguishes questions from statements in Swedish (Hadding-Koch & Studdert-Kennedy 1964), Venda (Ziervogel, Wentzel & Makuya 1972: 147), Kinyarwanda (Sibomana 1974: 185), English (Pierrehumbert 1980), Japanese (Pierrehumbert & Beckman 1988: 75), Kinande (Hyman 1990: 114), Chichewa (Myers 1996), and German (Féry 1993: 73). Questions are also marked by raising of the pitch range and/or reduction of phrasal pitch downtrends, as in Lingala (Guthrie 1940), Kongo (Carter 1973), Danish (Thorsen 1978), Kikuyu (Clements & Ford 1981), Hausa (Inkelas & Leben 1990), Jita (Downing 1995), and Kipare (Herman 1996). Final low tones mark yes–no questions in languages such as Chickasaw (Gordon 2005), Akan (Genzel & Kügler, 2020), and the languages surveyed in Rialland (2009). But the use of a high tone with phrase accent positioning has not been reported before for yes–no questions.

The previous descriptions of yes–no question intonation in Luganda were based on impressionistic transcriptions. The present study aims to clarify the difference between yes–no questions and statements in Luganda with an acoustic production experiment, comparing the two kinds of sentences across a range of lexical tone configurations.

## 2 Background on Luganda tone

There are three contrasting lexical tone categories in Luganda: high, low, and falling (Tucker 1962; Cole 1967; Stevick 1969a; Hyman 1982; Hyman & Katamba 1993, 2010). The minimal pair in (4) provides an example of the contrast between high and low tone (Snoxall 1967).

(4) a. [kùwólà]	b. [kùwòlà]
ku-wól-a	ku-wol-a
INF-grow.cold-FV	INF-lend-FV
‘to grow cold’	‘to lend’

The distribution of the three tone categories depends on syllable type. Luganda has a phonemic length contrast in both vowels and consonants, as exemplified by minimal pairs such as *kumala* ‘to finish’ – *kumaala* ‘to plaster’, and *kuba* ‘to be’ – *kubba* ‘to steal’ (Tucker 1962, Snoxall 1967, Clements 1986). Tucker (1962) defines the distribution of the tones in terms of a distinction between long and short syllables. A long syllable, according to Tucker, is one with a long vowel or a coda (i.e. the first half of a long consonant), while a short syllable is an open one with a short vowel. There is a contrast between high and low tone in both long and short syllables, but the contrast between falling and high tone only occurs in long syllables. Both high and falling tone are characterized by a rise in f<sub>0</sub> followed immediately by a fall, but they differ in that the f<sub>0</sub> peak occurs earlier in the syllable in falling tone than in high tone (Myers, Namyalo & Kiriggwajjo 2019).

A high tone in Luganda is subject to unbounded leftward tone spread. Within a tone phrase, as defined by Hyman, Katamba & Walusimbi (1987) and Pak (2008), a high tone extends leftward to the second syllable after a high-toned syllable, or, if there is no such

syllable, to the second syllable of the tone phrase (Hyman & Katamba 2010). The result of this process is an extended sequence of high-toned syllables, as in (5), where the high sequence is boldfaced. Spaces are omitted between the words in this and subsequent transcriptions in order to incorporate the long vowels and diphthongs that arise when a vowel at the end of one word is juxtaposed with a vowel at the beginning of the next word (Tucker 1962, Clements 1986, Myers 2020).

- (5) [ò mù lè : n z à : **mí r á n : á m ú n ò : n ò**]
- |            |                |             |         |
|------------|----------------|-------------|---------|
| O-mu-lenzi | a-mir-a        | nnamúnye    | ono     |
| IV-1-boy   | 3SG-swallow-FV | 1A.bird.sp. | 1A.this |
- ‘The boy is swallowing this bird.’

The only lexical high tone in this sentence belongs to the second syllable of the object noun *nnamúnye* ‘bird’. This extends leftward to the second syllable of the verb *amira* ‘he/she is swallowing’, which is the first word of the tone phrase that contains *nnamúnye*. The f<sub>0</sub> trajectory of such multi-syllable high-tone spans is described by Myers, Selkirk & Fainleib (2018).

Another high tone that is subject to this high tone spread is an intonational boundary tone H% found in statements. This high tone is manifested as a plateau of high-toned syllables extending from the sentence-final syllable leftward to the second syllable after a high-toned syllable, or the second syllable of the tone phrase. The final high-toned span is boldfaced in the example in as in (6).

- (6) [ò mù l w à : n à : **l é r é n : ó n í**]
- |                |            |           |
|----------------|------------|-----------|
| O-mu-lwany-i   | a-ler-a    | e-nnyoni  |
| IV-1-fight-NOM | 3SG-get-FV | IV-9.bird |
- ‘The fighter is carrying a bird.’

In this sentence, there are no words with a lexical high tone, yet there is a span of high-toned syllables extending from the final syllable of the sentence to the second syllable of the verb *alera* ‘he/she is carrying’, the first word of the final tone phrase. Hyman & Katamba (1993) attribute such spans to a boundary tone H%. Its meaning and distribution are unclear: Hyman (1982) describes it as characteristic of list intonation, while Hyman & Katamba (2010) describe it as indicating ‘finality’. The citation-form transcriptions in Cole (1967), Snoxall (1967), and Stevick (1969a) all include this intonational high tone for all items ending in two or more syllables without a lexical high tone. Hyman & Katamba (2010) describe the final H% in statements as optional, but Myers et al. (2018) report that it occurred in every statement in their sample of statements ending in final phrases lacking lexical high tones.

### 3 Experiment

The difference between yes-no questions and statements has been described in Luganda, but the description has been incomplete, and has the vagueness and subjectivity to be found in any generalizations about speech production based on impressionistic transcriptions. The present study reports the results of an acoustic production experiment in which yes-no questions and statements were compared across conditions differing in the location of the last lexical high tone.

Both the peak f<sub>0</sub> value and the timing of the f<sub>0</sub> peak were measured. In those sentences with a lexical high tone, Hyman’s (1982) description of the question-marking tone as ‘super-high’ leads to the expectation that the f<sub>0</sub> maximum would be greater in questions than in

**Table 1** Participants.

Subject	Gender	Age	Home district
S1	Male	58	Masaka
S2	Male	24	Kampala
S3	Female	24	Mubende
S4	Female	34	Kayunga
S5	Male	47	Wakiso
S6	Male	52	Mubende
S7	Male	57	Kampala
S8	Male	44	Butambala
S9	Female	79	Masaka
S10	Male	54	Masaka
S11	Female	34	Wakiso
S12	Male	45	Mubende
S13	Male	71	Kayunga
S14	Male	82	Luweero
S15	Female	60	Kampala
S16	Male	39	Masaka
S17	Male	38	Rakai
S18	Male	64	Kampala
S19	Female	62	Masaka

statements. Moreover, since this super-high tone is described as following the lexical high tone, we would hypothesize that the  $f_0$  peak will be later in questions than in statements in the case of sentences with a lexical high tone. The materials differed in the location of the last lexical high tone in order to test the claim that the location of the super-high intonational tone depends on the location of that lexical tone.

For sentences without lexical high tone, on the other hand, the description of Hyman (1982) would lead us to expect a local  $f_0$  peak on the second syllable of the sentence, at an  $f_0$  level higher than that for a lexical high tone in the same position. The later descriptions of Hyman (1990) and Hyman & Katamba (2011), on the other hand, would lead us to expect that the questions should have lower  $f_0$  than statements throughout the final interval. The description of Stevick (1969a: 27) predicts in particular that  $f_0$  at the end of the utterance will be lower in such questions than in corresponding statements.

### 3.1 Method

#### 3.1.1 Participants

Nineteen adult native speakers of Luganda participated in the study, ranging in age from 24 to 82. Six were female, and 13 were male. The relevant information for each one is listed in Table 1.

The participants came from all over the Central region of Uganda, where Luganda is spoken. Seventeen of them lived in the Kampala area at the time of the experiment, and two lived in the United States. They all grew up speaking Luganda, and at the time of the experiment spoke Luganda every day. They were also all fluent English speakers, English being a national language and *lingua franca* in Uganda.

Special effort was made to recruit participants from a broad range of age groups. The experimental descriptions of Luganda tone by Myers et al. (2019) and Myers et al. (2018) differed in important aspects from the description found in the previous literature (e.g. Tucker 1962, Cole 1967, Stevick 1969a), which could be due to a change in pronunciation between

speakers of that earlier time and contemporary speakers. Such a change would be evidenced by systematic differences in the measured properties between older and younger speakers, which could only be detected in a subject pool that varies sufficiently in age.

### 3.1.2 Materials

There were four classes of sentence in the study, depending on the position of the last lexical high tone in the sentence: HLL (lexical high tone on the antepenultimate syllable), LHL (lexical high tone on the penultimate syllable), LLH (lexical high tone on the final syllable), and LLL (no lexical high tone in the sentence). Examples of each class of sentence are provided in (7a–c) and all test sentences are listed in the appendix.

(7) a. HLL

Omulangira yalámula.

o-mu-langira y-a-lámul-a

IV-1-prince 3SG.SBJ-PST-judge-FV

‘The prince judged.’

b. LHL

Omulongo yalúma.

o-mu-longo y-a-lúm-a

IV-1-twin 3SG.SBJ-PST-bite-FV

‘The twin bit.’

c. LLH

Omulimi mulalú.

o-mu-lim-i mu-lalú

IV-1-cultivate-NMLZ 1-crazy

‘The farmer is crazy.’

d. LLL

Omulenzi alwanya omulimi.

o-mu-lenzi a-lwany-a o-mu-lim-i

IV-1-boy 3SG.SBJ-fight-FV IV-1-cultivate-NMLZ

‘The boy fights the farmer.’

In all sentence types, the last lexical high tone in the sentence was in a short syllable, and there were sonorant consonants preceding and following the vowel of that syllable. In none of the sentences did the lexical high tone meet the conditions for leftward spread (described in Section 2), so each lexical high tone was associated with just a single syllable. Half of the sentences produced were statements, and half were yes–no questions. The only orthographic difference between the yes–no question and the corresponding statement was the final punctuation mark: a question mark, or a period.

There were between four and six sentences in each class. All sentences were presented to the participants more than once in the recording session, yielding 15 tokens per participant

per condition. There were thus 15 tokens  $\times$  2 levels of Speech Act (Statement/Question)  $\times$  4 levels of Tone Position (HLL/LHL/LLH/LLL) = 120 tokens per speaker. There were 19 speakers, so a total of 2280 sentences in the study.

A total of 207 tokens were excluded from the analysis. All 120 tokens were excluded for participant S16, whose yes–no questions displayed no consistent pattern of question marking, differing in the number and location of  $f_0$  peaks even in different repetitions of the same question. For the other participants, 27 tokens were excluded because they included a pause within the measurement interval. Twenty-seven question tokens were excluded (all but one produced by S3), because they displayed an alternative question-marking strategy with an  $f_0$  peak on the sentence-final syllable, as in English. This might well represent a genuine alternative construction within Luganda, but the measurements for such tokens are not comparable to the ones with the usual Luganda question-marking pattern (which also made up the majority of questions for S3). Six tokens were excluded because the test word was replaced by another word, and 22 because the test word was produced with a lexical tone pattern violating the criteria for that sentence class. Two tokens had such reduced consonants that it was impossible to delimit the test syllable, and three had interruption of modal voicing in a critical part of the  $f_0$  trajectory of the test syllable. These exclusions left 2073 tokens for analysis.

### 3.1.3 Procedure

Seventeen of the participants were recorded at Makerere University in Kampala, Uganda, using a Shure SM10A head-mounted microphone and a Zoom H4n solid-state recorder, with a sampling rate of 44.1 kHz and 16-bit amplitude resolution. Two of the participants were recorded at their home in the United States, using the same microphone, sampling rate and amplitude resolution, but with a Marantz PMD 670 solid-state recorder.

The sentences were presented to the participants in a PowerPoint slideshow on a laptop computer, with each sentence on a separate slide. To avoid confusion, since the questions and statements differ orthographically only in the final punctuation, questions and statements were elicited in separate blocks. Within each block, the order of sentence stimuli was randomized. The sentences for this study were interspersed with sentences for other studies, which acted as distractors for this study.

Participants were instructed to read each sentence to themselves, and then to produce it without internal pauses, as a separate utterance (rather than as a member of a list). Participants were told that if they were not satisfied with their initial production, they could keep saying the sentence until they felt they had it right. They proceeded at their own pace through the sentences, but were instructed to finish saying a sentence before pressing the key to bring on the next one, and were asked to redo any sentence that was not preceded by a sufficient pause from the preceding sentence. When the participant produced a particular stimulus sentence more than once, the last one was selected for analysis, unless it had a clear internal pause or slip.

### 3.1.4 Measurements

Acoustic measurements were made using Praat (Boersma & Weenink 2013). For sentence types with a lexical high tone (HLL, LHL, LLH), the measurement interval extended from the onset of the syllable with the lexical high tone (S1) to the end of the immediately following syllable (S2), if there was one, and otherwise to the end of S1. The onset of each syllable was marked at the end of the amplitude drop from the preceding vowel. The offset of the utterance was marked at the end of voicing.

The duration of both S1 and S2 were measured.  $F_0$  measurements were made automatically using a script, with adjustments for each speaker for pitch range, voicing threshold, silence threshold, and octave jump cost. All  $f_0$  measurements were subjected to 10 Hz

smoothing. In sentence types with a lexical high tone, which were those with a localized  $f_0$  peak on a particular syllable, the following  $f_0$  measurements were made in the S1–S2 measurement domain:

- (8) a.  $F_0$  maximum: The maximum  $f_0$  within the measurement interval
- b. Offset  $f_0$ : The  $f_0$  at the end of voicing in the utterance
- c. Peak delay: The duration of the interval from the onset of S1 to the  $f_0$  maximum
- d. Relative peak delay: Peak delay divided by the duration of S1
- e.  $F_0$  sequence:  $F_0$  at each 10% increment of the duration of S1 and the duration of S2

Where there was no lexical high tone (LLL), there was no local peak corresponding to a lexical high tone, and the measurement interval was the whole utterance. The second syllable of the verb ( $S_{\text{medial}}$ ) was marked off, since that would be the onset of the final H% plateau in LLL statements, according to the usual pattern of tone spread (Hyman 1982). This divided the utterance into an initial interval leading up to this syllable, and a final interval extending from that syllable to the end of voicing in the utterance.  $F_0$  was sampled at every 20% increment within each of these two intervals.

### 3.1.5 Predicted differences

The primary goal of the experiment was to test the hypotheses about how Luganda yes–no questions differ in  $f_0$  from statements. In those sentences with a lexical high tone (HLL, LHL, LLH), the  $f_0$  maximum is expected to be greater in yes–no questions than in statements, reflecting the ‘super-high’ nature of the question-marking intonational tone. Relative peak delay is further expected to be greater in yes–no questions than in statements, reflecting the position of that intonational tone immediately following the lexical high tone.

In LLL sentences, there is no local  $f_0$  peak on a particular syllable. In statements, the final stretch of syllables extending from the sentence-final syllable to the second syllable of the verb are expected to have a higher  $f_0$  than the syllables preceding those in the sentence, due to the final H% span (Myers et al. 2018). In yes–no questions, on the other hand, the description of Hyman (1982) would lead us to expect a local  $f_0$  peak on the second syllable of the sentence, at an  $f_0$  level higher than that for a lexical high tone in the same position. The later descriptions of Hyman (1990) and Hyman & Katamba (2011), on the other hand, would lead us to expect that the questions should have lower  $f_0$  than statements throughout the final interval corresponding to the H% in statements. Following the description of Stevick (1969a: 27), offset  $f_0$  would be expected to be lower in LLL questions than in LLL statements. The literature has not made clear whether this difference in offset  $f_0$  between questions and statements extends to sentences with a lexical high tone.

The location of the last lexical high tone was varied in this experiment in order to test the distributional claim that the question-marking intonational tone is lodged immediately following that lexical tone. But the location of this high-toned syllable relative to the end of the phrase would also be expected to affect the measurements. For example, the  $f_0$  maximum might be lower for peaks closer to the end of the sentence, given the pervasiveness of down-trends in  $f_0$  values over the course of the phrase in languages (Lieberman & Pierrehumbert 1984, Poser 1984, Pierrehumbert & Beckman 1988).

Relative peak delay is also expected to be greater the farther that lexical high tone is from the end of the phrase. In other words, the  $f_0$  peak is expected to occur earlier in the syllable if that syllable is closer to the end of the phrase. Such a gradient effect of phrase-position on peak delay has been observed in Spanish (Prieto, van Santen & Hirschberg 1995) and Persian (Sadeghi 2017). Earlier  $f_0$  peaks for high tones in final compared to nonfinal syllables has been reported in English (Silverman & Pierrehumbert, 1990), Spanish (Prieto et al. 1995), Palermo Italian (Grice 1995), Chichewa (Myers 1999), Kinyarwanda (Myers 2003),

Moroccan Arabic (Yeou 2004), Serbian (Smiljanić 2006), German (Mücke & Hermes 2007), and Chickasaw (Gordon 2008).

Because relative peak delay is derived by dividing peak delay by syllable duration, a difference in relative peak delay according to sentence position could be due to differences in either of these component measurements. A lower relative peak delay in one condition can be attained by a shorter  $f_0$  rise, reflected in peak delay. Or it could be due to a longer syllable, reflected in S1 duration. These measurements will be examined to unpack any effects on relative peak delay.

With respect to S1 duration, phrase-final segments are generally longer than comparable phrase-medial ones (Klatt 1975). This pattern of final lengthening has been found in languages all over the world (Myers & Hansen 2007). It is gradient, in the sense that the effect is greater the closer the relevant segment is to the end of the phrase (Lindblom, Lyberg & Holmgren 1981, Turk 1999), reflecting a gradual deceleration of articulatory movements as the speaker approaches pause (Edwards, Beckman & Fletcher 1991). From these considerations, one might expect that S1 would have a greater duration the closer it was to the end of the phrase (in this case, the end of the sentence).

### 3.1.6 Statistical analysis

Mixed linear regression models were fit to the data using the packages *lme4* (Bates et al. 2014) and *lmerTest* (Kuznetsova, Brockhoff & Christensen 2014) in R (R Core Team 2017). For analyses of the sentences with a lexical high tone, the fixed effects were Speech Act (Question/Statement), and Tone Position (the number of syllables separating the lexical high-toned syllable from the end of the phrase: HLL = 2, LHL = 1, LLH = 0). Pairwise comparison of the three levels in Tone Position were performed using the Tukey method in the *emmeans* package (Lenth et al. 2020). For the analysis of LLL, the only fixed effect was Speech Act. Random intercepts were included for Speaker and for Item (sentence), and random slopes for the interaction of Speaker with fixed effects. If the model failed to converge, the analysis was re-run with the random slopes omitted one by one until convergence was attained. The alpha level was  $p < .05$ .

## 3.2 Results

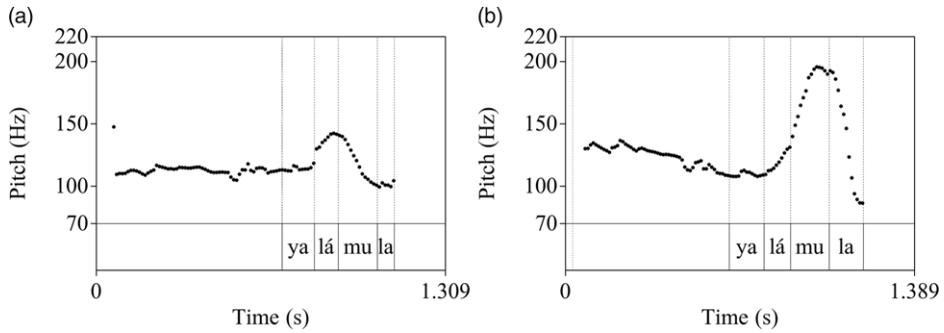
Figures 1–4 present sample annotated pitch tracks for representative sentences from Speaker 6. Syllables in the test word are marked off by vertical lines in the pitch track. For glosses and morpheme breakdowns of the examples, see the appendix.

In Figures 1 and 2, displaying sentences with a nonfinal lexical high tone, the statement has an  $f_0$  rise beginning near the onset of the high-toned syllable and ending in an  $f_0$  peak near the end of that syllable. In the corresponding questions, the  $f_0$  rise that starts at about the same point as in the statement, but continues on to a peak in the syllable following that one. This  $f_0$  peak is higher than in the corresponding statement.

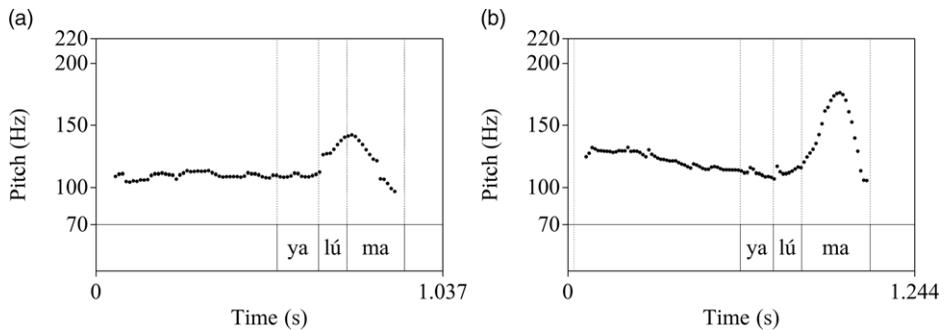
In Figure 3, the lexical high tone is in the final syllable of the sentence. The  $f_0$  peak is higher and occurs later in the syllable in the question than in the statement. The test syllable is also longer than in the HLL and LHL conditions.

In Figure 4, there is no lexical high tone, and no single-syllable  $f_0$  peak in either the statement or the question. The annotation marks the three words of the sentence, and also the three syllables of the medial verb (*alima*). In the statement in Figure 4a, the final H% plateau extends from the second syllable of the verb to the end of the sentence. In the question in Figure 4b, on the other hand,  $f_0$  at the start of the sentence is higher than in the statement, and there is a gradual  $f_0$  fall to the end of the sentence.

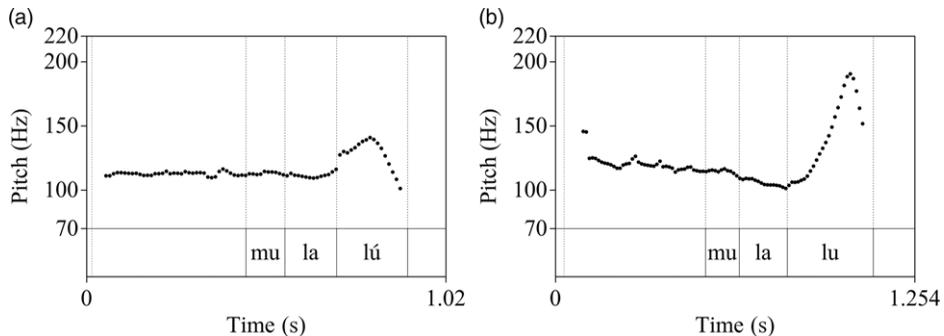
Pooling across speakers, Figure 5 presents the mean  $f_0$  sequence within the measurement interval for each condition with a lexical high tone.  $F_0$  measurements were made at each 10% increment of the duration of the target S1 syllable, and the same for S2 (the following syllable), if there was one. In order to allow the pooling of measurements from participants



**Figure 1** Sample pitch tracks of HLL sentences. (a) Statement: *Omulangira yalámula*. 'The prince judged'. (b) Question: *Omulangira yalámula?* 'Did the prince judge?'.

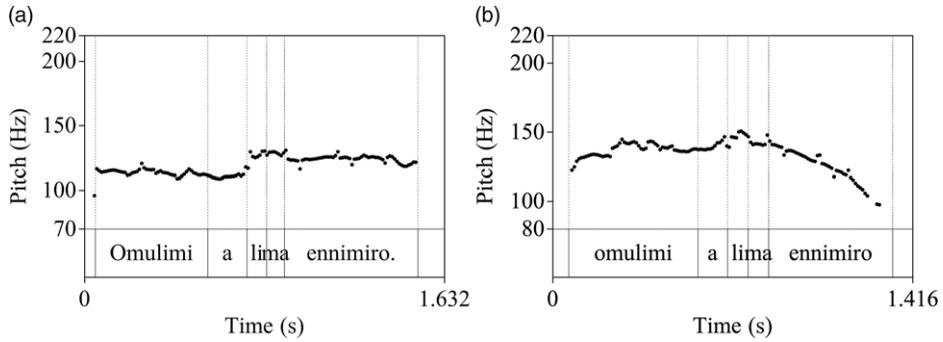


**Figure 2** Sample pitch tracks of LHL sentences. (a) Statement: *Omulongo yalúma*. 'The twin bit'. (b) Question: *Omulongo yalúma?* 'Did the twin bite?'.

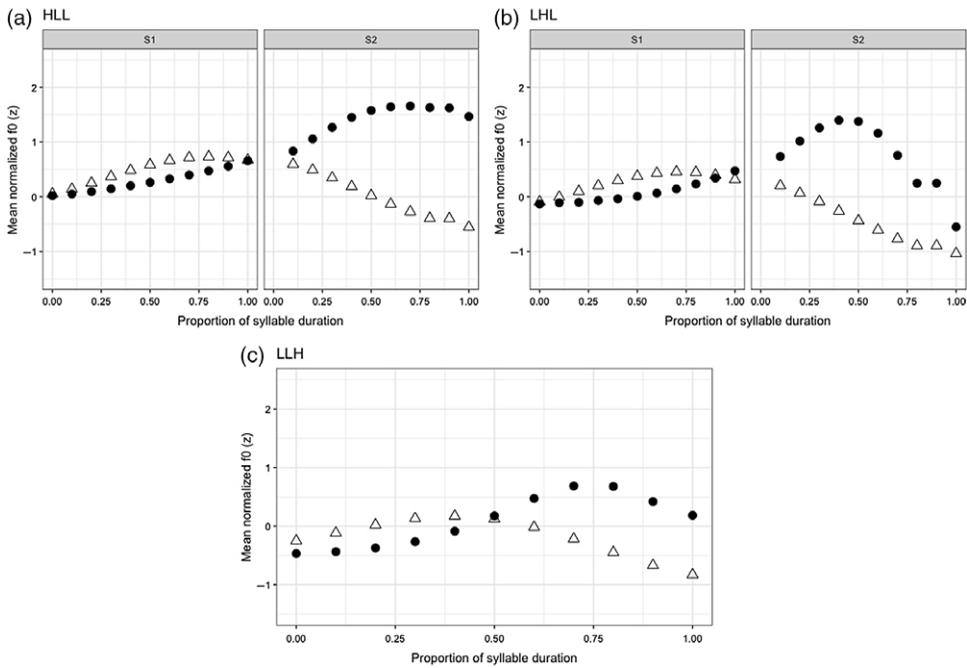


**Figure 3** Sample pitch tracks of LLH sentences. (a) Statement: *Omulimi mulalú*. 'The farmer is crazy'. (b) Question: *Omulimi mulalú?* 'Is the farmer crazy?'.

with quite different pitch ranges,  $f_0$  measurements have been normalized relative to the mean and standard deviation of the individual speaker, so that each plotting point represents the average for that timepoint of the normalized  $f_0$  in z-scores. Questions are marked by filled circles and statements by hollow triangles. In Figures 5a–b, the first panel shows S1, and the second panel shows S2. In Figure 5c, there is just one panel, displaying the sentence-final syllable with the lexical high tone. These normalized plots show that in the conditions with a lexical high tone, the question has a later and a higher  $f_0$  peak than the corresponding statement.



**Figure 4** Sample pitch tracks of LLL sentences. (a) Statement: *Omulimi alima ennimiro*. 'The farmer cultivates a garden'. (b) Question: *Omulimi alima ennimiro?* 'Is the farmer cultivating a garden?'.



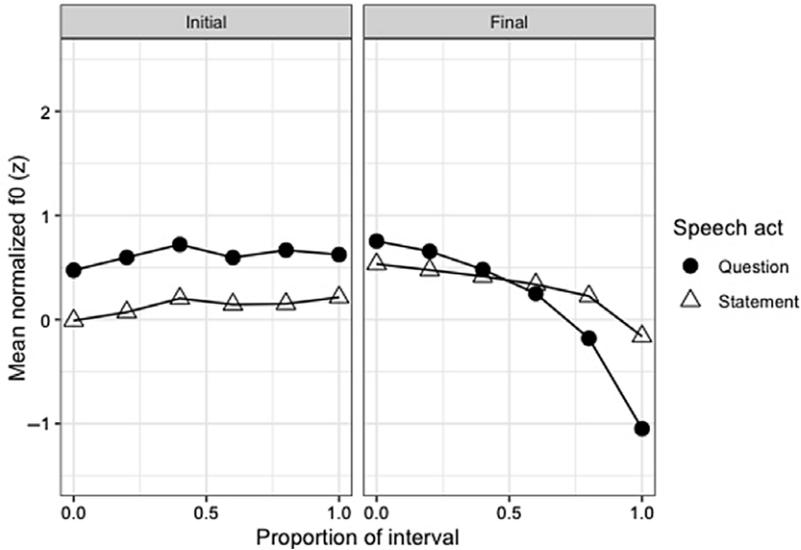
**Figure 5** Mean normalized  $f_0$  (z) by proportion of measurement interval: (a) HLL, (b) LHL, (c) LLH. Hollow triangles mark measurement points in statements, and filled circles mark those in questions.

Figure 6 presents the normalized  $f_0$  trajectory for LLL sentences. Since there are no local  $f_0$  peaks in these sentences without lexical high tones, the whole sentence is represented in two intervals: the initial interval from the utterance onset to the onset of the second syllable of the verb, and a final interval from the offset of that syllable to the offset of the utterance.  $f_0$  was measured at the onset and offset of each interval, and at each 20% increment of the interval duration.

The normalized  $f_0$  is higher throughout the initial interval in questions than in statements. In statements,  $f_0$  rises from the onset of  $S_{\text{medial}}$  (at the end of the first panel) to the offset of that same syllable (at the beginning of the second panel), representing the beginning of the final H% plateau. In questions, on the other hand, normalized  $f_0$  begins to fall at the offset of  $S_{\text{medial}}$ , ending at an offset  $f_0$  value lower than that for statements.

**Table 2** Model of Maximum f0. Maximum f0  $\sim$  Speech Act  $\times$  Tone Position + (1 + Speech Act | Participant) + (1 | Item).

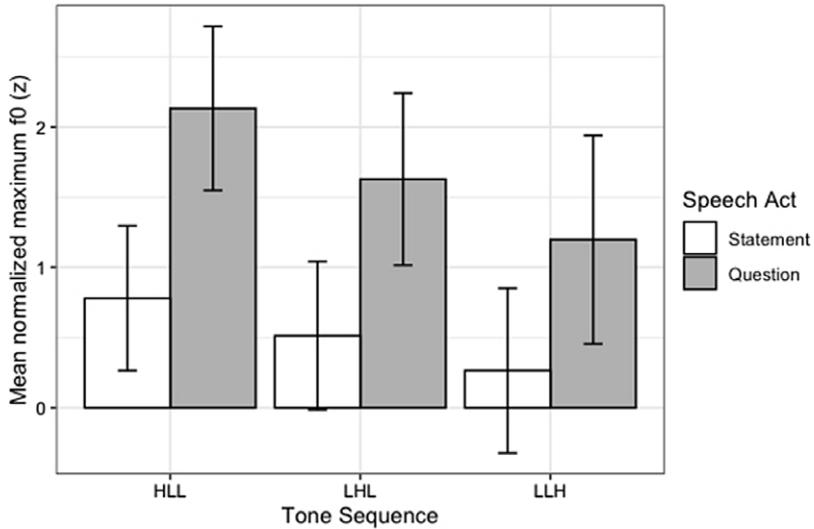
Fixed effect	$\beta$	<i>df.</i>	<i>t</i>	<i>p</i>
<b>(Intercept)</b>	<b>157.61</b>	<b>21.5</b>	<b>14.5</b>	<b>&lt; .001</b>
<b>Speech Act</b>	<b>31.73</b>	<b>39.8</b>	<b>4.2</b>	<b>&lt; .001</b>
<b>Tone Position</b>	<b>9.88</b>	<b>27.8</b>	<b>3.7</b>	<b>.001</b>
Speech Act $\times$ Tone Position	7.39	27.2	1.9	.06

**Figure 6** Mean normalized f0 (z) by proportion of measurement interval in the Initial and the Final intervals for LLL sentences.

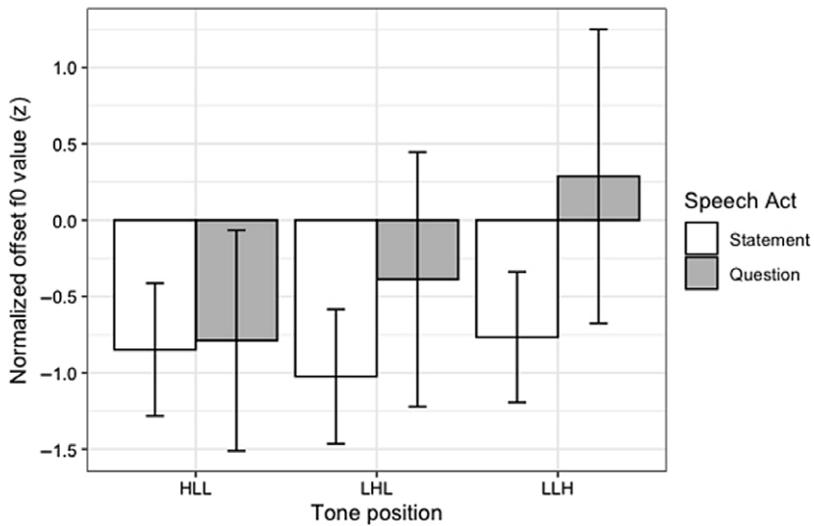
### 3.2.1 Sentences with a lexical high tone (HLL, LHL, LLH)

Figure 7 presents the mean normalized f0 maximum in sentences with a lexical high tone, broken down by Tone Position and Speech Act. The mean normalized f0 maximum was greater in questions (1.65) than in statements (0.52), and greater in lexical high tones that were farther from the end of the sentence: HLL (1.45), LHL (1.07), LLH (0.73). The pattern held across participants, all of whom had a higher mean normalized f0 maximum in questions than in statements. Except for S3, all participants also had the same descending pattern for position: HLL > LHL > LLH. For S3, the mean normalized f0 maximum was higher for HLL than for LHL, but LHL had a lower mean than LLH.

A model of maximum f0 is presented in Table 2. Here the dependent variable is unnormalized maximum f0 (Hz), rather than the normalized values (z) depicted in Figure 7, since inclusion of Participant as a random effect takes into account the variation among participants in mean f0 level. Significant effects ( $p < .05$ ) are highlighted in this and subsequent tables by boldface. Both the main effects of Speech Act and Tone Position were significant, and there was no significant interaction between them. The coefficient for Speech Act is 31.73, indicating that the predicted value for questions (the marked level) is 31.73 Hz greater than that for statements (the default level), when all other effects are factored out. The coefficient for Tone Position is 9.88, indicating that each syllable that separates the last lexical high tone from the end of the sentence increases the predicted f0 maximum by 9.88 Hz. The three tone position classes were compared pairwise to each other, and all pairs were significantly different.



**Figure 7** Mean normalized maximum f0 (z) by Tone Position and Speech Act. (Error bars represent standard deviation.)



**Figure 8** Mean normalized offset f0 (z) by Tone Position and Speech Act. (Error bars represent standard deviation.)

A reviewer points out that three of the sentences (in the LHL and LLH conditions) include a lexical high tone preceding the test high tone, and suggests that the effect of Tone Position could plausibly be due to downstep triggered by the preceding high tone in these items. However, Tone Position still has a significant effect in the same direction if these sentences are excluded.

The offset f0 was the f0 value at the end of the utterance. The group means for normalized offset f0 are presented in Figure 8.

Overall, mean normalized offset f0 was higher in questions (-0.29) than in statements (-0.88), and it was higher in positions closer to the end of the sentence than in those farther from the end: HLL (-0.82), LHL (-0.71), LLH (-0.24). The difference between question

**Table 3** Model of Offset f0. Offset f0  $\sim$  Speech Act  $\times$  Tone Position + (1 + Speech Act | Participant) + (1 | Item).

Fixed effect	$\beta$	<i>d.f.</i>	<i>t</i>	<i>p</i>
<b>(Intercept)</b>	<b>123.11</b>	<b>18.7</b>	<b>13.3</b>	<b>&lt; .001</b>
<b>Speech Act</b>	<b>36.95</b>	<b>29.1</b>	<b>6.9</b>	<b>&lt; .001</b>
Tone Position	-0.70	28.2	-0.4	.66
<b>Speech Act <math>\times</math> Tone Position</b>	<b>-17.13</b>	<b>28.1</b>	<b>-7.6</b>	<b>&lt; .001</b>

**Table 4** Models of Offset f0, broken down by Tone Position subset. Offset f0  $\sim$  Speech Act + (1 | Participant) + (1 | Item).

Tone Position subset	$\beta$	<i>d.f.</i>	<i>t</i>	<i>p</i>
HLL	1.79	499.0	1.1	.26
<b>LHL</b>	<b>21.68</b>	<b>7.5</b>	<b>8.1</b>	<b>&lt; .001</b>
<b>LLH</b>	<b>35.85</b>	<b>6.7</b>	<b>9.4</b>	<b>&lt; .001</b>

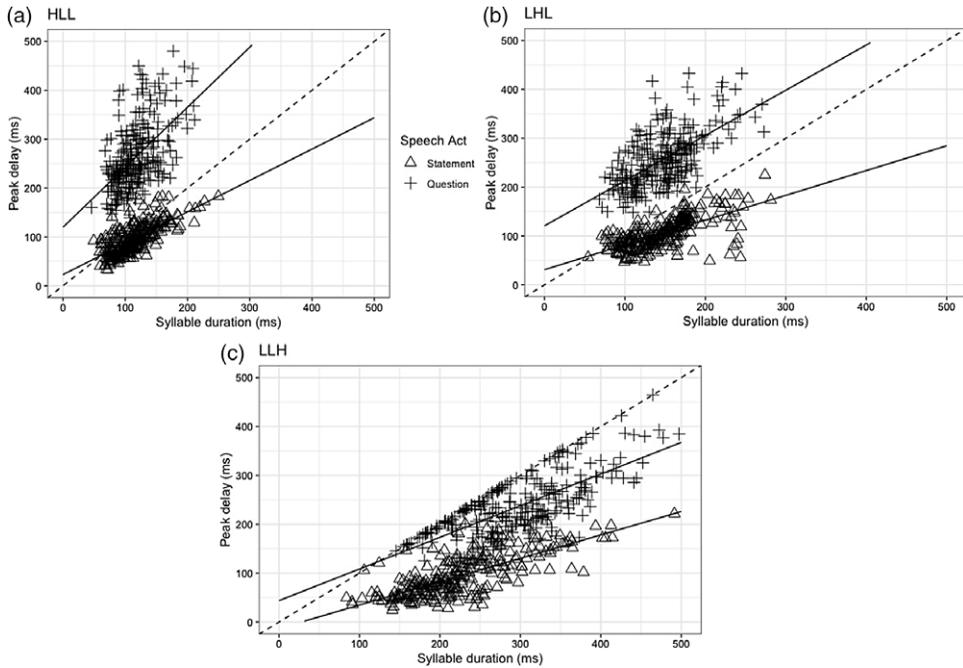
**Table 5** Models of Offset f0, broken down by Speech Act subset. Offset f0  $\sim$  Tone Position + (1 | Participant) + (1 | Item).

Speech Act subset	$\beta$	<i>d.f.</i>	<i>t</i>	<i>p</i>
Statement	-0.56	13.3	-0.4	.70
<b>Question</b>	<b>-17.96</b>	<b>14.8</b>	<b>-10.1</b>	<b>&lt; .001</b>

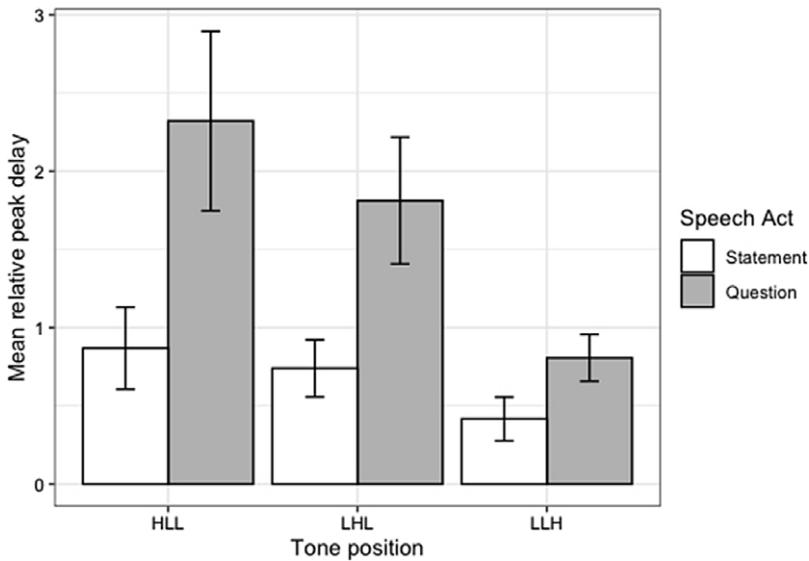
and statement is greater in LLH than in LHL, and greater in LHL than in HLL. These effects on offset f0 are modeled in Table 3. There is a main effect of Speech Act, and an interaction of Speech Act with Tone Position. The factor Speech Act has a positive coefficient in the model (36.95), while the interaction has a negative coefficient (-17.13). This indicates that questions (the marked level of Speech Act) have a significantly higher offset f0 than statements, but this effect is reduced with each syllable that separates the lexical high tone from the end of the sentence.

To explore the interaction, each Tone Position subset was submitted to an analysis with the same structure as that in Table 3, but without the Tone Position factor. The results for the Speech Act variable are given in Table 4. The analysis is then broken down by Speech Act class in Table 5. Offset f0 was significantly greater in questions than in statements in Tone Position classes LLH and LHL, but not in HLL. On the other hand, the effect of Tone Position on offset f0 was limited to questions.

Peak delay for each token in the sample is plotted in Figure 9 against S1 duration, broken down by Tone Position and Speech Act. The dashed line marks  $x=y$ , so points on that line would mark instances in which the f0 peak is exactly at the end of the syllable. Points below that line mark f0 peaks within S1, while points above that line mark f0 peaks in the next syllable. The two solid lines are the regression lines through the question and statement points. Both lines slope upward in all three graphs, indicating that longer S1 duration is associated with longer peak delay. The crosses marking questions are generally above the triangles marking statements, reflecting the fact that the f0 peak delay in questions (mean = 252 ms) was generally greater than in statements (mean = 99 ms). The statement markers are generally below the  $x=y$  line, indicating that in statements the peak is within S1, while the question markers in HLL and LHL lie above that line, indicating that in these cases the peak for questions is in the syllable following S1. There is no such syllable in the case of LLH, so in that case the question points are above the statement points, but within S1.



**Figure 9** Peak delay (ms) by S1 duration (ms), Tone Position and Speech Act: (a) HLL, (b) LHL, (c) LLH.



**Figure 10** Mean relative peak delay by Tone Position and Speech Act. (Error bars represent standard deviation.)

Relative peak delay provides a measure of how the peak is timed with respect to the syllable, since it gives the proportion of S1 duration at which the  $f_0$  peak is attained. If the value is below 1, the  $f_0$  peak lies within S1, and if it is over 1, the peak lies in the following syllable. The group means for relative peak delay are presented in Figure 10. Across tone positions, mean relative peak delay was higher in questions (1.64) than in statements (0.67). It was also higher in lexical high-tone positions that are farther from the end of the sentence:

**Table 6** Model of Relative Peak Delay. Relative peak delay  $\sim$  Speech Act  $\times$  Tone Position + (1 + Speech Act + Tone Position | Participant) + (1 | Item).

Fixed effect	$\beta$	<i>df.</i>	<i>t</i>	<i>p</i>
<b>Intercept</b>	<b>0.48</b>	<b>27.0</b>	<b>4.9</b>	<b>&lt; .001</b>
<b>Speech Act</b>	<b>0.49</b>	<b>30.5</b>	<b>3.4</b>	<b>.002</b>
<b>Tone Position</b>	<b>0.17</b>	<b>31.9</b>	<b>2.4</b>	<b>.02</b>
<b>Speech Act <math>\times</math> Tone Position</b>	<b>0.57</b>	<b>24.4</b>	<b>5.9</b>	<b>&lt; .001</b>

**Table 7** Model of Relative Peak Delay, broken down by Tone Position subset. Relative peak delay  $\sim$  Speech Act + (1 | Participant) + (1 | Item).

Tone Position subset	$\beta$	<i>df.</i>	<i>t</i>	<i>p</i>
<b>HLL</b>	<b>1.46</b>	<b>19.3</b>	<b>8.8</b>	<b>&lt; .001</b>
<b>LHL</b>	<b>1.17</b>	<b>9.7</b>	<b>8.1</b>	<b>&lt; .001</b>
<b>LLH</b>	<b>0.38</b>	<b>10.5</b>	<b>6.4</b>	<b>&lt; .001</b>

HLL (1.59) > LHL (1.41) > LLH (0.61). The difference between questions and statements was greater in the earlier tone positions than in the later ones. Both the differences due to Speech Act and those due to Tone Position held for each participant considered separately.

A model of relative peak delay is presented in Table 6. Both main effects are significant, as well as the interaction. In a pairwise test of the three levels in Tone Position, all pairs were found to be significantly different.

To explore the interaction, we examine the three Tone Position subsets. Each subset was submitted to an analysis with the same structure as that in Table 6, but without the Tone Position factor. The results for the Speech Act factor for each Tone Position subset are given in Table 7.

The peak is significantly later in questions than in corresponding statements in all three Tone Position classes, but the effect of Speech Act is greater for Tone Position classes with the lexical high tone farther from the end of the sentence.

However, the interpretation of these results is complicated by the fact that relative peak delay is a derived measurement calculated by dividing peak delay by S1 duration. The effects of Speech Act and Tone Position on relative peak delay, as seen in Tables 6 and 7, could therefore be due to effects on peak delay or S1 duration. We therefore examine these measurements next.

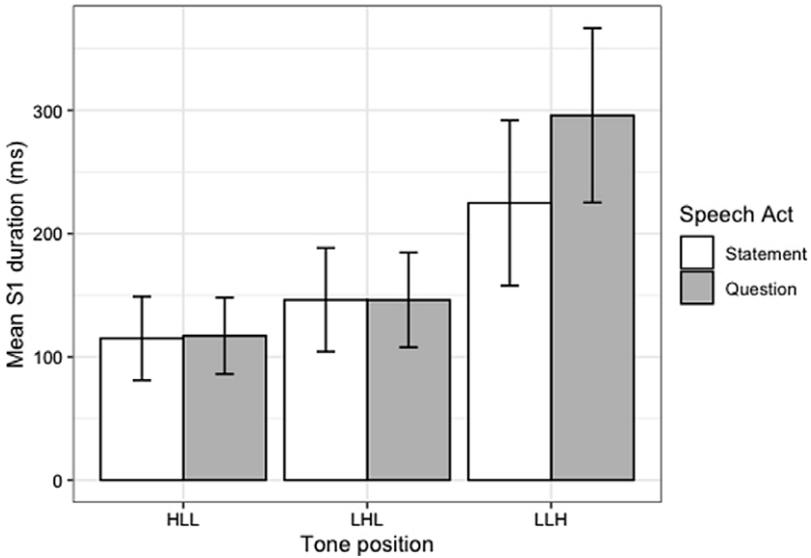
The mean duration for the high-toned syllable S1 (ms) is presented broken down by Speech Act and Tone Position in Figure 11. Mean S1 duration was greater in later sentence positions than in earlier ones: LLH (260 ms) > LHL (146 ms) > HLL (116 ms). It was also greater in questions (187 ms) than in statements (162 ms), but this effect was clearly limited mainly to the LLH position class.

The analysis of S1 duration is presented in Table 8. There is a main effect of Tone Position, and a significant interaction of Speech Act with Tone Position. To investigate the interaction, Tone Position subsets were each submitted to an analysis with the same structure as that in Table 8, but without the Tone Position factor. Speech Act did not have a significant effect in any of these models. In the pairwise comparison of the Tone Position classes, S1 duration was significantly greater in LLH than in HLL and LHL, but there was no difference between HLL and LHL. The results thus provide evidence that the test syllable was longer when it was phrase-final than when it was phrase medial.

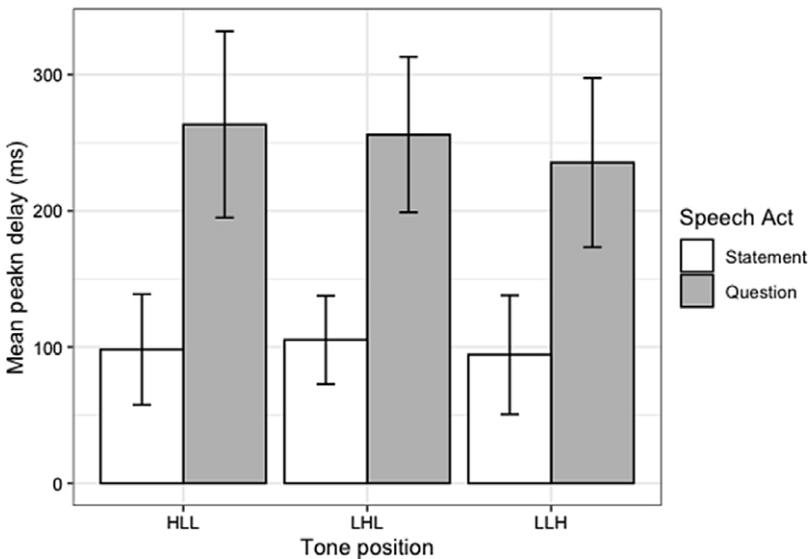
Group means for peak delay, the other component of relative peak delay, are presented in Figure 12. Mean peak delay was greater in questions (252 ms) than in statements (99 ms), but it did not vary greatly according to Tone Position: HLL (181 ms), LHL (180 ms), LLH (165 ms).

**Table 8** Model of S1 duration.  $S1\ duration \sim \text{Speech Act} * \text{Tone Position} + (1 + \text{Speech Act} | \text{Participant}) + (1 | \text{Item})$ .

Fixed effect	$\beta$	<i>d.f.</i>	<i>t</i>	<i>p</i>
<b>Intercept</b>	<b>216.64</b>	<b>27.7</b>	<b>11.6</b>	<b>&lt; .001</b>
Speech Act	51.39	26.5	12.0	.06
<b>Tone Position</b>	<b>-49.54</b>	<b>28.6</b>	<b>-3.9</b>	<b>&lt; .001</b>
<b>Speech Act × Tone Position</b>	<b>-38.38</b>	<b>26.0</b>	<b>-2.1</b>	<b>.046</b>



**Figure 11** Mean S1 duration by Tone Position and Speech Act. (Error bars represent standard deviation.)



**Figure 12** Mean peak delay by Tone Position and Speech Act. (Error bars represent standard deviation.)

**Table 9** Model of Peak Delay. Peak delay  $\sim$  Speech Act  $\times$  Tone Position + (1 + Speech Act + Tone Position | Participant) + (1 | Item).

Fixed effect	$\beta$	<i>df.</i>	<i>t</i>	<i>p</i>
<b>Intercept</b>	<b>102.69</b>	<b>30.5</b>	<b>7.6</b>	<b>&lt; .001</b>
<b>Speech Act</b>	<b>141.32</b>	<b>38.0</b>	<b>6.8</b>	<b>&lt; .001</b>
Tone Position	-1.12	32.0	-0.1	.91
Speech Act $\times$ Tone Position	10.27	27.3	0.8	.44

**Table 10** Models of F0 in LLL, broken down by timepoint. F0 at timepoint  $\sim$  Speech Act + (1 + Speech Act | Speaker) + (1 | Item).

Interval	Position	Random slope	$\beta$	<i>df</i>	<i>t</i>	<i>p</i>
Initial	<b>0%</b>	<b>Yes</b>	<b>16.89</b>	<b>16.2</b>	<b>6.4</b>	<b>&lt; .001</b>
	<b>20%</b>	<b>Yes</b>	<b>18.17</b>	<b>16.8</b>	<b>7.2</b>	<b>&lt; .001</b>
	<b>40%</b>	<b>Yes</b>	<b>17.65</b>	<b>16.9</b>	<b>6.4</b>	<b>&lt; .001</b>
	<b>60%</b>	<b>Yes</b>	<b>15.28</b>	<b>16.6</b>	<b>5.3</b>	<b>&lt; .001</b>
	<b>80%</b>	<b>Yes</b>	<b>17.48</b>	<b>16.8</b>	<b>5.7</b>	<b>&lt; .001</b>
	<b>100%</b>	<b>Yes</b>	<b>13.75</b>	<b>16.9</b>	<b>4.5</b>	<b>&lt; .001</b>
Final	<b>0%</b>	<b>No</b>	<b>7.33</b>	<b>481.7</b>	<b>6.0</b>	<b>&lt; .001</b>
	<b>20%</b>	<b>No</b>	<b>5.9</b>	<b>483.1</b>	<b>4.8</b>	<b>&lt; .001</b>
	40%	No	1.94	481.0	1.6	.11
	<b>60%</b>	<b>No</b>	<b>-3.3</b>	<b>482.1</b>	<b>-2.8</b>	<b>.006</b>
	<b>80%</b>	<b>No</b>	<b>-13.77</b>	<b>477.1</b>	<b>-10.8</b>	<b>&lt; .001</b>
	<b>100%</b>	<b>No</b>	<b>-30.62</b>	<b>467.3</b>	<b>-22.2</b>	<b>&lt; .001</b>

The analysis is presented in Table 9. The main effect of Speech Act is significant, but not the effect of Tone Position, or the interaction.

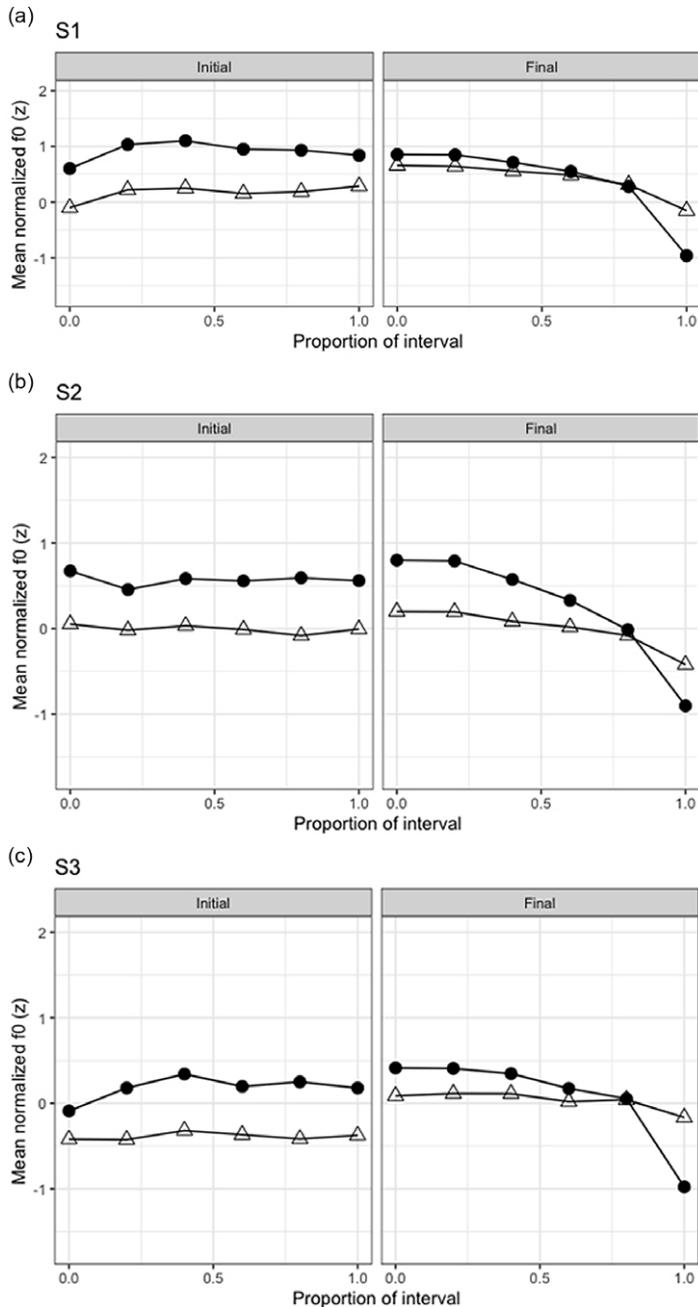
The difference in relative peak delay between questions and statements (Tables 6 and 7) was reflected in a parallel difference in peak delay (Table 9), and so was due to the longer absolute peak delay in questions than in statements. On the other hand, the fact that relative peak delay was greater in earlier tone positions than in later ones was reflected in a parallel trend in syllable duration (Table 8), and so was due to the fact that phrase-final syllables were longer.

### 3.2.2 Sentences without a lexical high tone (LLL)

In the LLL sentences without a lexical high tone, there is no local f0 maximum to measure. Instead, f0 was sampled at 20% increments of the interval up to S<sub>medial</sub> and the interval from the offset of that syllable to the end of the utterance. The f0 values for questions and statements were compared at each measurement position, and the results are presented in Table 10. The column 'Random slope' indicates whether or not the random slope factor Participant  $\times$  Speech Act was included in the final model.

Questions had a significantly higher f0 value at each measurement point up through the one 20% through the interval from the end of S<sub>medial</sub> to the end of the utterance. The next measurement point after that had no significant difference between questions and statements, and then in the last three measurement points, the f0 value in questions was significantly lower than in statements.

The LLL sentences were by far the most variable sentence type in their f0 production, and participants produced in particular the yes-no questions in this sentence type with greater hesitancy and more attempts than in any other condition. Figure 13 presents



**Figure 13** Mean normalized  $f_0(z)$  by proportion of measurement interval in the Initial and the Final intervals for LLL sentences for each speaker. Hollow triangles mark measurement points in statements, and black circles mark those in questions.

the time-normalized  $f_0$  trajectories for LLL questions and statements for each individual participant (solid circles marking questions and hollow triangles marking statements).

For 14 of the 18 participants,  $f_0$  in questions was higher than in statements throughout the initial interval (in the first panel). This wasn't the case for S9, S11, S13, or S14. Furthermore, 14 of 18 participants had a high plateau in questions extending from the beginning of the

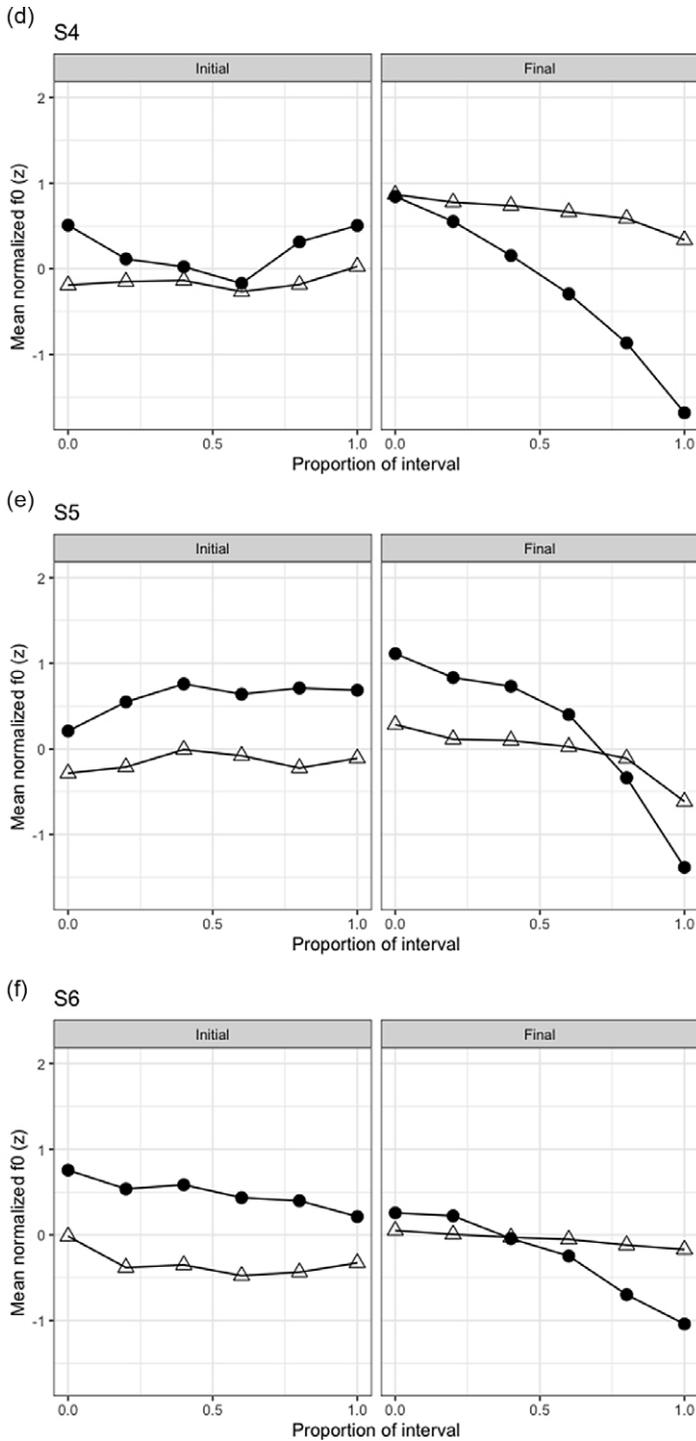
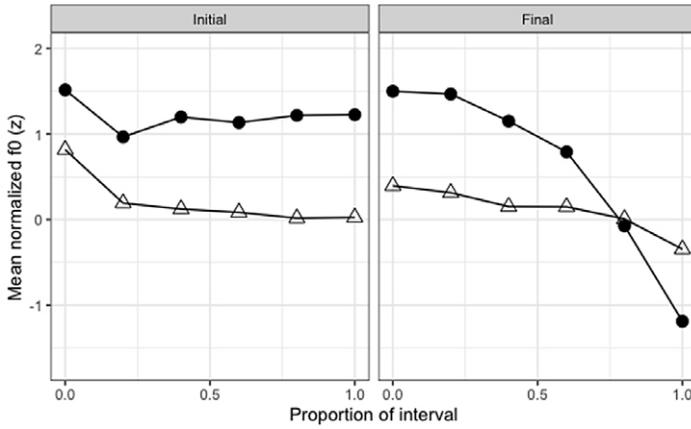
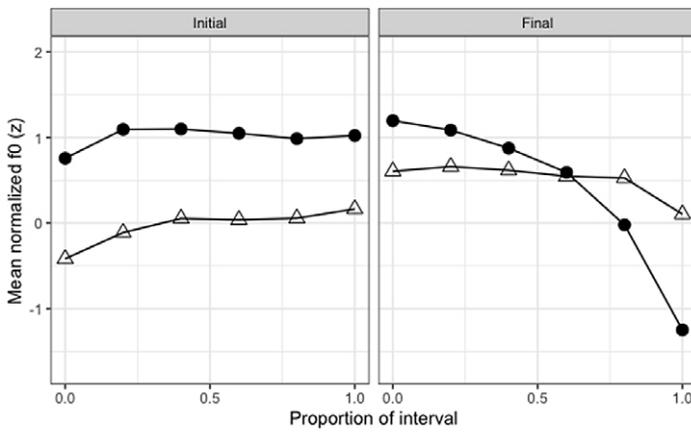


Figure 13 Continued.

(g) S7



(h) S8



(i) S9

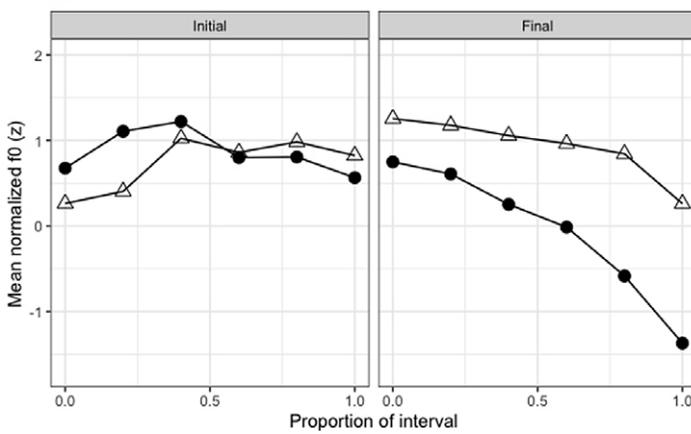


Figure 13 Continued.

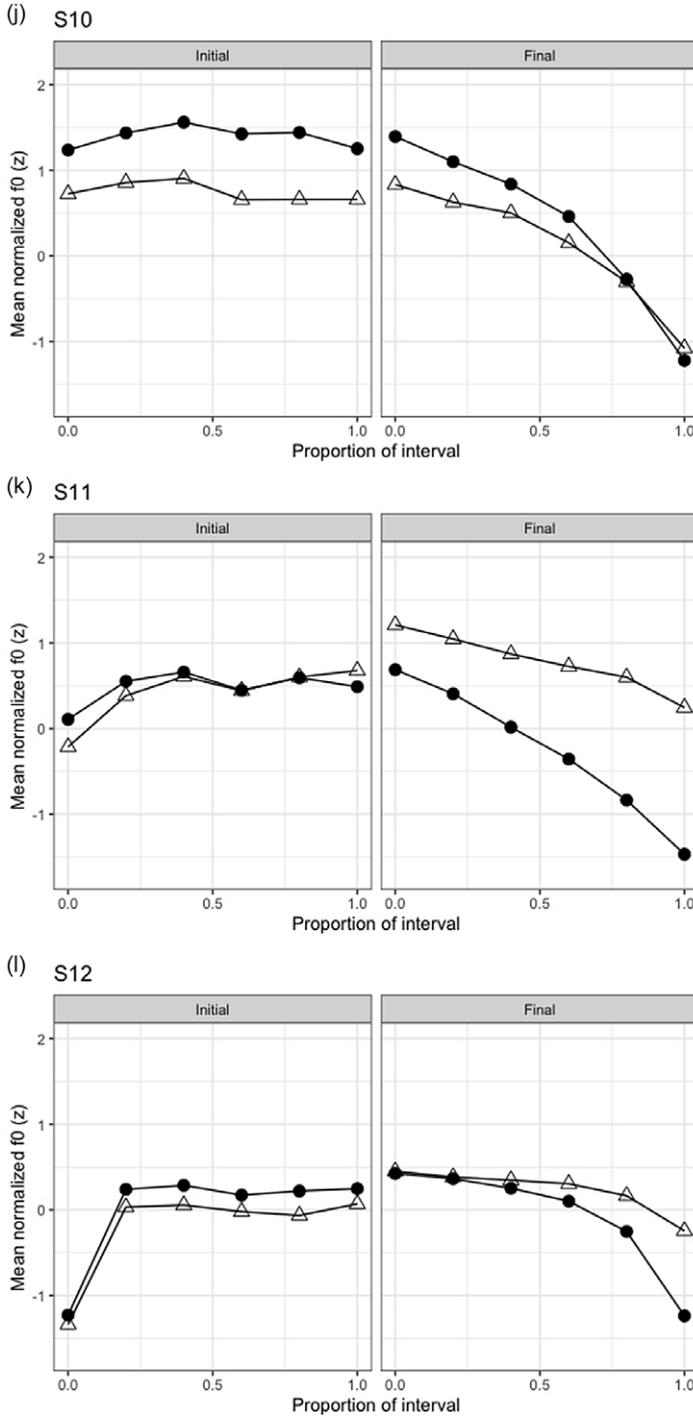


Figure 13 Continued.

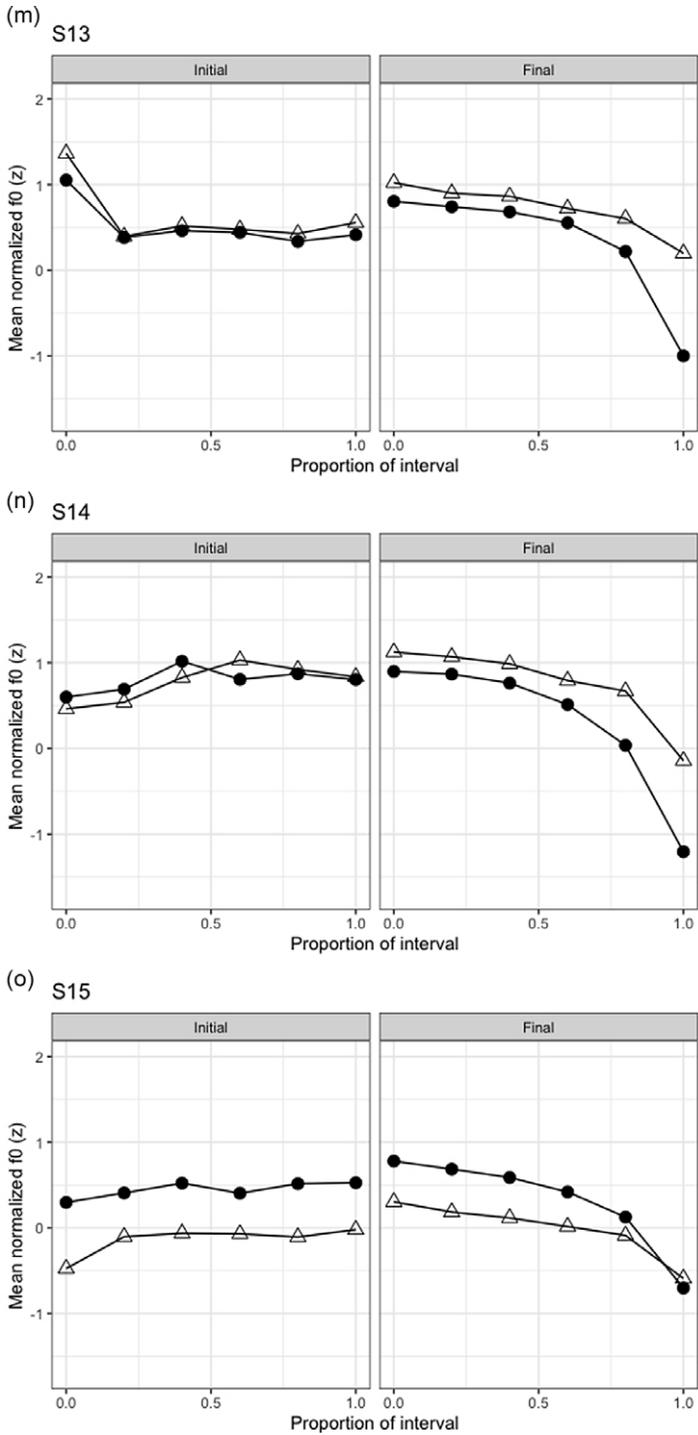


Figure 13 Continued.

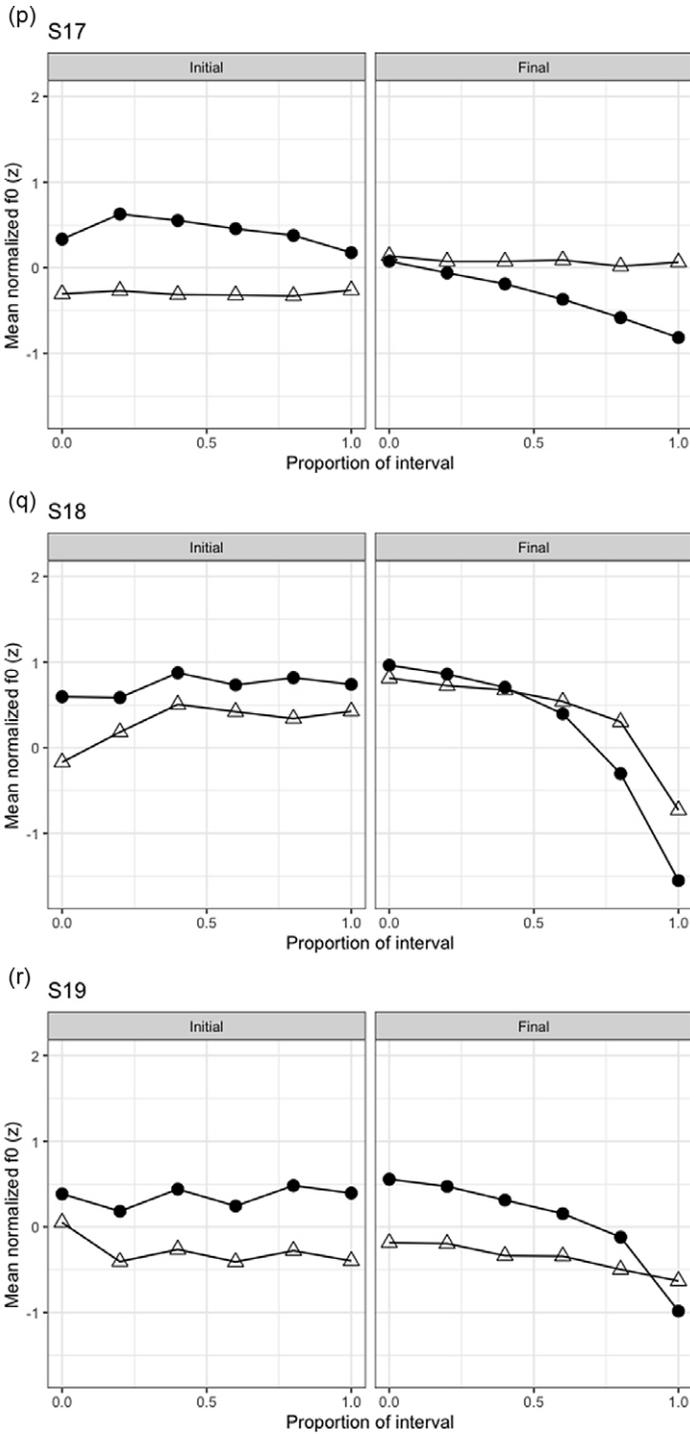


Figure 13 Continued.

sentence to the first syllable of the second interval, with the final  $f_0$  fall beginning in the final interval. But 4 of the participants had a peak early in the first interval in questions, followed by a steady fall to the end (S1, S6, S9, S17). All participants had an increase in  $f_0$  in statements from the end of the first interval to the beginning of the second, on the second syllable of the verb, but participants differ in how large this increase was, and in the slope of the subsequent trajectory in the second interval. The final  $f_0$  measurement point was lower in questions than in statements for all participants, but the difference between them was very small for S10 and S15. None of these patterns of variation among the participants coincided with groups defined by gender, home district, or age (Table 1).

## 4 Discussion

### 4.1 The difference between yes–no questions and statements in Luganda

This study has clarified how yes–no questions in Luganda differ from corresponding statements. It has done so partly because it is the first study of the topic based on objective acoustic measurements, but also because it included cases that have not been covered in the quite brief discussions of the matter in the previous literature.

In sentences with a lexical high tone, Hyman (1982) described the yes–no question as differing from the corresponding statement in having a ‘super-high’ tone immediately following the last lexical high tone. We expressed this quantitatively in the hypothesis that the yes–no question would have an  $f_0$  peak that was significantly higher and later than in the corresponding statement. This hypothesis was supported in our study. In both the question and the statement, the  $f_0$  rise began near the beginning of the S1 syllable (the one with the last lexical high tone in the sentence), and it rose in both speech act types throughout most of that syllable. In the question, however,  $f_0$  continued to rise during the syllable following S1, attaining an  $f_0$  peak in that syllable that was higher than would be found in a lexical high tone in a comparable position. Both the  $f_0$  maximum and the relative peak delay were greater in the yes–no question than in the corresponding statement.

In sentence types in which there was a syllable following the lexical high-tone syllable, i.e. in HLL and LHL, the super-high  $f_0$  peak marking yes–no question occurred in that following syllable. But in LLH, in which the final lexical high tone occurred in the sentence-final syllable, the question-marking peak occurred in that same final syllable, but later in the syllable than the peak in the corresponding statements. This is consistent with how the pattern is described by Hyman (1982, 1990) and Stevick (1969a), though examples of the LLH case were not provided in any of those works.

All of the examples in Hyman (1982, 1990) and Hyman & Katamba (2011) have just one lexical high tone, but Hyman (1990: 122) specifies that the yes–no question marker occurs after the *last* lexical high tone. The materials in the current experiment included examples in which there was more than one lexical high tone, and these confirmed that this generalization was correct. In the sentences in (9), for example, there are lexical high tones on the penultimate syllables of both the verb *yanóna* and the object *omulére*, and the yes–no question in (9b) has the super-high tone only after the second of those.

(9) a. [òmùsòmìjànónò:mùlérè]

Omusomi yanóna omulére.

o-mu-som-i                      y-a-nón-a                      o-mu-lére

IV-1-read-NMLZ              3SG.SBJ-PST-pick.up-FV              IV-3-flute

‘The student fetched the flute.’

- b. [òmùsòmijànónò:mùléré]  
 Omusomi yanóna omulére?  
 o-mu-som-i          y-a-nón-a                  o-mu-lére  
 IV-1-read-NMLZ      3SG.SBJ-PST-pick.up-FV    IV-3-flute.Q  
 ‘Did the student fetch the flute?’

In all cases in this sample with more than one lexical high tone, the question-marking super-high tone occurred only after the last one.

Stevick (1969a: 27) stated that the question-marking pitch rise only occurred when the sentence-final word had a lexical high tone. Hyman (1982, 1990) did not restrict the pattern to the final word, but only cited one-word examples. It was therefore unclear from these descriptions what the intonation of a yes–no question would be if the last lexical high tone wasn’t in the final word of the sentence. As it happens, all the sentences in this experiment that have a lexical high tone have it in the final word of the sentence. But in the preliminary pilot work leading up to this experiment, there were sentences such as those in (10), in which the last lexical high tone was in a nonfinal word.

- (10) a. [òmùlè:nzìjàmánpò:mù:ntùòmùlàlà]  
 Omulenzi yamánya omuntu omulala.  
 o-mu-lenzi    y-a-mány-a                  o-mu-ntu      o-mu-lala  
 IV-1-boy      3SG.SBJ-PST-know-FV    IV-1-person    IV-1-different  
 ‘The boy knew a different person.’
- b. [òmùlè:nzìjàmánpó:mù:ntùòmùlàlà]  
 Omulenzi yamánya omuntu omulala?  
 o-mu-lenzi    y-a-mány-a                  o-mu-ntu      o-mu-lala  
 IV-1-boy      3SG.SBJ-PST-know-FV    IV-1-person    IV-1-different\Q  
 ‘Did the boy know a different person?’

Here the last lexical high tone in the statement in (10a) is on the second syllable of the verb, *yamánya* ‘he/she knew’, which is the third word from the end of the sentence. In the corresponding question in (10b), the three speakers who produced these test items consistently put the super-high tone on the syllable following the lexical high tone in the verb. This confirms the generalization of Hyman (1990) that it is the last high tone that is relevant, and provides evidence against the generalization of Stevick (1969a) that the relevant high tone must be in the last word of the sentence.

In sentences without a lexical high tone (LLL), Hyman (1982) described the yes–no question as having a super-high tone on the second syllable. This description would be supported if there was a peak on the second syllable of the yes–no question with a greater  $f_0$  maximum than in the corresponding syllable of the statement. As it turned out, there was no one-syllable  $f_0$  maximum anywhere in either the question or the statement in these sentences, so this description was not supported in this study.

On the other hand, Stevick (1969a), Hyman (1990) and Hyman & Katamba (2011) all described LLL questions as having low tone throughout with lower final pitch than in the corresponding statement, while the statement had a final high plateau throughout the final phrase. These descriptions would be supported if  $f_0$  was low throughout the initial interval in both questions and statements, and then higher in statements than in questions for the final interval (reflecting the final H% in statements).

This description was only partially supported by our findings. Most participants (14 out of 18) had higher  $f_0$  in questions than in statements throughout the initial interval of the sentence from the beginning to the onset of  $S_{\text{medial}}$ , the second syllable of the verb. Generalizing across speakers,  $f_0$  was significantly higher in questions than in statements at each measurement point throughout this interval. This was not expected based on any of the previous descriptions, and it suggests the presence of an intonational high tone in this interval of the yes–no questions that is absent in the corresponding statements.

On the other hand, the  $f_0$  trajectories in the second interval, starting with  $S_{\text{medial}}$ , were more in line with these descriptions. All speakers showed an  $f_0$  rise in the LLL statements over the course of the syllable  $S_{\text{medial}}$ , though they varied in whether that syllable was followed by a plateau or a gradual decline.  $F_0$  was significantly lower in the question than in the statement at all measurement points in the final 40% of this interval, with the greatest differences in the utterance-final measurement point. The difference in offset  $f_0$  is consistent with the accounts of Stevick (1969a) and Hyman (1990), who described this as an important difference between statements and yes–no questions in sentences without lexical high tones.

## 4.2 Effects of Tone Position

In sentences with a lexical high tone, the position of the final lexical high tone was systematically varied, in order to test whether that would affect the position of the intonational tone marking yes–no intonation. The  $f_0$  maximum in yes–no questions occurred immediately following the position of the final  $f_0$  peak in the corresponding statement, supporting the description of Hyman (1982, 1990), according to which the super-high tone marking yes–no questions is associated immediately following the last lexical high tone in the sentence.

This variation in the position of the final lexical high tone also had other measurable effects. Maximum  $f_0$  was greater the farther that lexical high tone was from the end of the phrase: HLL > LHL > LLH. Such an effect reflects the general downtrend in  $f_0$  values over the course of the phrase, and in particular the effects of  $f_0$  lowering in phrase-final position (Lieberman & Pierrehumbert 1984, Poser 1984, Pierrehumbert & Beckman 1988). The effect was gradient in Luganda, in that a high tone in the antepenultimate syllable had a significantly higher maximum  $f_0$  value than one in the penultimate syllable, and the latter in turn had a significantly higher maximum  $f_0$  value than one in the final syllable.

The effect of Tone Position on offset  $f_0$ , however, went in the opposite direction from the lowering effects seen with maximum  $f_0$ . In questions,  $f_0$  at the end of the utterance was higher when the lexical high tone was closer to the end: LLH > LHL > HLL. This can be interpreted as a coarticulatory effect of the high  $f_0$  peak in questions. It takes time for  $f_0$  to return to baseline values after such a high peak, and the less time there is for this recovery from the peak, the higher  $f_0$  will be when time runs out at the end of the utterance.

Tone Position also had an effect on the timing of the  $f_0$  peak. Relative peak delay was greater the farther that lexical high tone was from the end of the phrase. In other words, the  $f_0$  peak occurred earlier in the syllable if that syllable was closer to the end of the phrase. Such a gradient effect of phrase-position on  $f_0$  timing has been observed in Spanish (Prieto et al. 1995) and Persian (Sadeghi 2017). The difference in Luganda was not due to differences in the duration of the  $f_0$  rise, but instead was associated with longer S1 duration in the final syllable of the phrase compared to nonfinal syllables (final lengthening). Peak delay is in general greater if syllable duration is greater, as we saw in Figure 9, but the additional syllable duration due to phrase position does not seem to count for this relation, leading to proportionally earlier  $f_0$  peaks in lengthened syllables, as found in English by Silverman & Pierrehumbert (1990).

## 4.3 Representations

The  $f_0$  patterns described above reflect sequences of tone categories, both lexical and intonational, and a context-sensitive system of phonetic implementation mapping those sequences

to f0 trajectories (Pierrehumbert 1980, Beckman & Pierrehumbert 1988). In this section we will consider the question of what kinds of tone categories could lead to the observed f0 patterns.

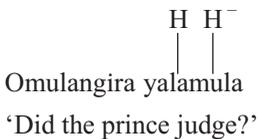
A tone category is associated with a syllable or, in the case of tone spread, a sequence of syllables. It is reflected in the f0 trajectory by a movement in f0 that is timed with respect to that syllable or sequence of syllables. Lexical tones belong to particular lexical items, and whether a lexical tone occurs in the utterance depends on which lexical items are there.

It is assumed here that the contrast between high tone and low tone in Luganda is a privative one between the presence of a high f0 target (H) and the absence of such a target (Stevick 1969b, Myers 1998). H tone in Luganda is realized with an f0 rise from the default low f0 level to a relatively high value, followed by a return to the default level (Myers et al. 2019), as in the realization of H\* in English (Pierrehumbert 1980). The realization of the falling tone of Luganda, which was not included in the current study, is the same as the high tone, except with an earlier f0 peak (Myers et al. 2019).

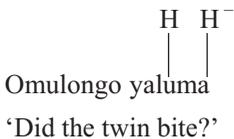
The final H% in statements is a boundary tone which is associated with the final syllable in the sentence (Hyman 1990). It is an intonational tone because it does not belong to any of the words that make up the sentence, and whether it occurs or not depends on what kind of sentence it is. This intonational high tone, like a lexical high tone, is subject to unbounded spread, as described above in Section 2, which spreads the tone leftward up to the syllable after a high-toned syllable, or the syllable after the onset of the phrase (Hyman & Katamba 2010). Hyman & Katamba (2010: 71) describes this boundary tone as optional in statements, but in our sample it occurred reliably in any statement that ended in a sequence of three or more lexically toneless syllables. As shown by Myers et al. (2018), the statement H% is realized with a lower f0 maximum than a comparable lexical high tone.

Yes–no questions in Luganda must be marked by an intonational high tone, since they differ from the corresponding statements in having an additional interval of raised f0, whether following a lexical high tone, or forming a peak or plateau at the beginning of the sentence. This is not a boundary H%, since it is not constrained to the final syllable of the intonational phrase. Rather, its positioning is parallel to the H<sup>-</sup> of English, which Pierrehumbert (1980) describes as occurring immediately following the nuclear pitch accent in the sentence. The question phrase accent H<sup>-</sup> occurs immediately following the final lexical H, i.e. as early as it can occur without preceding such a tone, as in (11a–c). If there is no lexical high tone, it occurs on the second syllable in the domain, either the sentence, as in (11d), or the tone phrase, as in (11e).

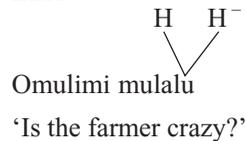
(11) a. HLL



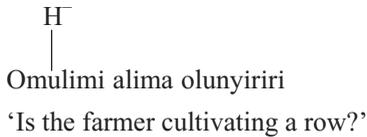
b. LHL



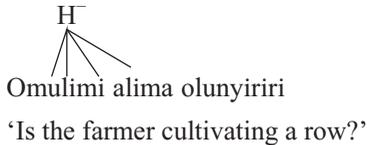
c. LLH



d. LLL



e. LLL



In LLL sentences, there is no lexical H tone to block the leftward path of  $H^-$ . For those speakers with an early  $f_0$  peak in LLL questions, the  $H^-$  is associated with the second syllable in the sentence, as in (11d). For those speakers with an initial  $f_0$  plateau, the  $H^-$  is instead associated with the second syllable of the final tone phrase, as defined by Hyman et al. (1987), which in these sentences corresponds to the final verb phrase (the verb and the following complement or modifier). The  $H^-$  is blocked from landing on the first syllable of the phrase, just as the leftward spread of a high tone within that tone phrase would be blocked from that first syllable (Hyman 1982). From that docking site, the  $H^-$  is subject to unbounded leftward spread to the second syllable of the sentence, as described by Hyman & Katamba (2010). The result is the representation in (11e).

The  $H^-$  is phonetically interpreted like other high tones in the language, except that it is assigned a higher  $f_0$  value when it occurs immediately following another high tone. According to the model in Table 2, the  $f_0$  peak in a question with a lexical high tone is 31.7 Hz higher than the corresponding statement, all else being equal. This upstep effect is similar to the raising of  $H\%$  after  $H^-$  in English polar questions (Pierrehumbert 1980), or the raising of intonational H after  $H^*$  in circumflex question intonation in Spanish (Torreira & Grice 2018).

The offset  $f_0$  value of a LLL yes–no question is lower than in a corresponding statement. This cannot be due to the  $H^-$  in questions, which is not at the end of the sentence in LLL sentences. Instead, it can be attributed to the fact that LLL statements end in a  $H\%$  plateau. There is no such high boundary tone in the corresponding questions, so they end at a lower  $f_0$  level.

According to this description, Luganda yes–no questions occupy a previously unattested place in intonational typology. These questions are marked by an intonational high tone, as is the case in many languages, but it is a mobile phrase accent, not a boundary tone restricted to the final syllable of the phrase, as in English (Pierrehumbert 1980), Japanese (Pierrehumbert & Beckman 1988) or Chichewa (Myers 1996). This phrase accent is positioned in the zone between the last pitch prominence in the phrase and the end of the phrase, as proposed for other languages by Pierrehumbert (1980) and Grice et al. (2000), but in Luganda that last pitch prominence is a lexical tone, not associated with a stress prominence.

The distribution of the  $H^-$  tone after the last marked element and otherwise at the beginning of the domain is reminiscent of parallel patterns of distribution in nonlinear phonology. For example, stress falls on the last heavy syllable and otherwise the first syllable in Eastern Chhemis or Huasteco (Hayes 1981). In Chaha, the 3rd masculine singular objective in perfectives is marked by labialization of the last labializable consonant in the base (McCarthy 1983). In Japanese mimetics, the ‘uncontrolled’ variant is marked by palatalization of the last coronal in the form and otherwise the initial consonant (Mester & Itô 1989). In each of these cases, an entity is positioned as close to one end as it can get without crossing a designated obstacle. The result is a pattern of distribution in which the entity occurs at the leftmost/rightmost obstacle, and in the absence of obstacles, at the rightmost/leftmost end.

## 5 Conclusion

In this paper, experimental acoustic evidence has been provided to test descriptive claims about the yes–no question intonation in Luganda. The results support a model in which the yes–no question is marked by an intonational high tone that is positioned immediately after the last lexical high tone, if there is one, and otherwise on the second syllable of the domain (tone phrase or sentence). In its positioning after the last tone target in the phrase, this intonational tone is parallel to the H<sup>-</sup> phrase accent posited in analyses of the intonational systems of European languages (Bruce 1977, Pierrehumbert 1980, Grice et al. 2000).

## Acknowledgements

The author would like to thank Dr. Saudah Namyalo, Sam and Rose Musoke, Anatole Kiriggwajjo, and Paul Bbosa for their invaluable help in designing and running this experiment, the 19 participants for sharing their knowledge of Luganda, and the editor, associate editor and reviewers for the feedback that helped get this article into shape.

## Appendix. Test sentences

In the following list of test sentences, the statement form of each statement is given. The representation is the standard orthography, augmented with acute accents marking lexical high tones. The question is identical in the orthography except with a question mark at the end of the sentence. The number in parentheses after the gloss indicates the number of repetitions for that sentence for each speaker for each sentence type.

### (A1) HLL

- a. Omulangira yalámula.  
 o-mu-langira                    y-a-lámul-a  
 IV-1-prince    3SG.SBJ-PST-judge-FV  
 ‘The prince judged.’ (3)
- b. Omulimi alírima.  
 o-mu-lim-i    a-lí-rim-a  
 IV-1-cultivate-NMLZ    3SG.SBJ-FUT-cultivate-FV  
 ‘The farmer will cultivate.’ (3)
- c. Omuntu yamúluma.  
 o-mu-ntu            y-a-mú-lum-a  
 IV-1-person    3SG.SBJ-PST-3SG.OBJ-bite-FV  
 ‘The person bit him/her.’ (3)
- d. Omufumbi yamúnona.  
 o-mu-fumb-i                    y-a-mú-non-a  
 IV-1-cook-NMLZ            3SG.SBJ -PST-3SG.OBJ-pick.up-FV  
 ‘The cook picked him/her up.’ (2)

- e. Omusomi alímalá.  
 o-mu-som-i                      a-lí-mal-a  
 IV-1-read-NMLZ              3SG.SBJ-FUT-finish-FV  
 ‘The student will finish.’ (2)
- f. Omulenzi alírwana.  
 o-mu-lenzi      a-lí-rwan-a  
 IV-1-boy              3SG.SBJ-FUT-fight-FV  
 ‘The boy will fight.’ (2)

## (A2) LHL

- a. Omulongo yalúma.  
 o-mu-longo    y-a-lúm-a  
 IV-1-twin              3SG.SBJ-PST-bite-FV  
 ‘The twin bit.’ (3)
- b. Omuntu yamíra.  
 o-mu-ntu              y-a-mír-a  
 IV-1-person              3SG.SBJ-PST-swallow-FV  
 ‘The person swallowed.’ (3)
- c. Omulangira yaláma  
 o-mu-langira    y-a-lám-a  
 IV-1-prince              3SG.SBJ-PST-be.in.good.health-FV  
 ‘The prince survived.’ (3)
- d. Omuntu munéne.  
 o-mu-ntu    mu-néne  
 IV-1-person 1-big  
 ‘The person is big.’ (3)
- e. Omusomi yanóna omulére.  
 o-mu-som-i              y-a-nón-a                      o-mu-lére  
 IV-1-read-NMLZ              3SG.SBJ-PST-pick.up-FV              IV-3-flute  
 ‘The student fetched the flute.’ (3)

## (A3) LLH

- a. Omulimi mulalú.  
 o-mu-lim-i                      mu-lalú  
 IV-1-cultivate-NMLZ    1-crazy  
 ‘The farmer is crazy.’ (6)

- b. Omuntu ono muwemú.

o-mu-ntu    ono    mu-wemú  
 IV-1-person    this/1    1-shameless  
 ‘This person is shameless.’ (3)

- c. Omulenzi yamíra omunyá.

o-mu-lenzi    y-a-mír-a                            o-mu-nyá  
 IV-1-boy    3SG.SBJ-PST-swallow-FV    IV-3-lizard  
 ‘The boy swallowed a lizard.’ (3)

- d. Omufumbi yamánya omulí.

o-mu-fumb-i    y-a-mány-a                            o-mu-l-í  
 IV-1-cook-NMLZ    3SG.SBJ-PST-know-FV    IV-1-eat-NMLZ  
 ‘The cook knew the eater.’ (3)

(A4) LLL

- a. Omulenzi alwanya omulimi.

o-mu-lenzi    a-lwany-a                            o-mu-lim-i  
 IV-1-boy    3SG.SBJ-fight-FV    IV-1-cultivate-NMLZ  
 ‘The boy fights the farmer.’ (3)

- b. Omulimi alima olunyiriri.

o-mu-lim-i                            a-lim-a                            o-lu-nyiriri  
 IV-1-cultivate-NMLZ    3SG.SBJ-cultivate-FV    IV-11-row  
 ‘The farmer cultivates a row.’ (3)

- c. Omulimi alima ennimiro.

o-mu-lim-i                            a-lim-a                            e-n-nim-iro  
 IV-1-cultivate-NMLZ    3SG.SBJ-cultivate-FV    IV-9-cultivate-NMLZ  
 ‘The farmer cultivates a garden.’ (3)

- d. Omulangira alamula omulimi.

o-mu-langira    a-lamul-a                            o-mu-lim-i  
 IV-1-prince    3SG.SBJ-judge-FV    IV-1-cultivate-NMLZ  
 ‘The prince judges the farmer.’ (3)

- e. Omuntu amala omulimu.

o-mu-ntu    a-mal-a                            o-mu-limu  
 IV-1-person    3SG.SBJ-finish-FV    IV-3-work  
 ‘The person finishes work.’ (3)

## Supplementary material

To view supplementary material for this article (including audio files to accompany the language examples), please visit <https://doi.org/10.1017/S0025100321000025>.

## References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2014. *Lme4*: Linear mixed-effects models using eigen and S4. <http://www.R-project.org>.
- Beckman, Mary & Julia Hirschberg. 1994. The ToBI annotation conventions. <http://www.speech.cs.cmu.edu/tobi/ToBI.6.html>.
- Boersma, Paul & David Weenink. 2013. Praat: Doing phonetics by computer (version Stanford, CA, 5.3.53). <http://www.praat.org/>.
- Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. Lund: CWK Gleerup.
- Carter, Hazel. 1973. *Syntax and tone in Kongo*. London: School of Oriental and African Studies (SOAS).
- Clements, G. N. 1986. Compensatory lengthening and consonant gemination in LuGanda. In Leo Wetzels & Engin Sezer (eds.), *Studies in compensatory lengthening*, 39–77. Dordrecht: Foris.
- Clements, G. N. & Kevin Ford. 1981. On the phonological status of downstep in Kikuyu. In Didier Goyvaerts (ed.), *Phonology in the 1980's [sic]*, 309–357. Amsterdam: John Benjamins.
- Cole, D. T. 1967. *Some features of Ganda linguistic structure*. Johannesburg: Witwatersrand University Press.
- Downing, Laura. 1995. The metrical domain of register raising in Jita: An optimality approach. In Francis Katamba (ed.), *Bantu phonology and morphology*, 28–39. Munich: Lincom Europa.
- Edwards, Jan, Mary E. Beckman & Janet Fletcher. 1991. The articulatory kinematics of final lengthening. *The Journal of the Acoustical Society of America* 89, 369–382.
- Féry, Caroline. 1993. *German intonational patterns*. Tübingen: Max Niemeyer Verlag.
- Genzel, Susanne & Frank Kügler. 2020. Production and perception of question prosody in Akan. *Journal of the International Phonetic Association* 50(1), 61–92.
- Gordon, Matthew. 2005. Intonational phonology of Chickasaw. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 301–330. Oxford: Oxford University Press.
- Gordon, Matthew. 2008. Pitch accent timing and scaling in Chickasaw. *Journal of Phonetics* 36, 521–535.
- Grice, Martine. 1995. *The intonation of interrogation in Palermo Italian*. Tübingen: Niemeyer.
- Grice, Martine, D. Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17, 143–185.
- Guthrie, Malcolm. 1940. Tone ranges in a two-tone language (Lingala). *Bulletin of the School of Oriental and African Studies* 10, 469–478.
- Hadding-Koch, Kerstin & Michael Studdert-Kennedy. 1964. An experimental study of some intonation contours. *Phonetica* 11, 175–185.
- Hayes, Bruce. 1981. *A metrical theory of stress rules*. Bloomington, IN: Indiana University Linguistics Club.
- Herman, Rebecca. 1996. Final lowering in Kipare. *Phonology* 13, 171–196.
- Hyman, Larry M. 1982. Globality and the accentual analysis of Luganda tone. *Journal of Linguistic Research* 2, 1–40.
- Hyman, Larry M. 1990. Boundary tonology and the prosodic hierarchy. In Sharon Inkelas & Draga Zec (eds.), *The phonology–syntax connection*, 109–125. Chicago, IL: The University of Chicago Press.
- Hyman, Larry M. & Francis X. Katamba. 1993. A new approach to tone in Luganda. *Language* 69, 34–67.
- Hyman, Larry M. & Francis X. Katamba. 2010. Tone, syntax and prosodic domains in Luganda. In Laura Downing, Annie Rialland, Jean-Marc Beltzung, Sophie Manus, Cédric Patin & Kristina Riedel (eds.), *Papers from the Workshop on Bantu Relative Clauses*, 69–98. Berlin: Zentrum für Allgemeine Sprachwissenschaft.
- Hyman, Larry M. & Francis X. Katamba. 2011. The tonology of WH questions in Luganda. *ZAS Papers in Linguistics* 55, 65–81.

- Hyman, Larry M., Francis X. Katamba & Livingstone Walusimbi. 1987. Luganda and the strict layer hypothesis. *Phonology Yearbook* 4, 87–108.
- Hyman, Larry M. & Kemmonyne C. Monaka. 2011. Tonal and non-tonal intonation in Shekgalagari. In Sonia Frota, Gorka Elordieta & Pilar Prieto (eds.), *Prosodic categories: Production, perception and comprehension*, 267–290. Dordrecht: Springer.
- Inkelas, Sharon & William Leben. 1990. Where phonology and phonetics intersect: The case of Hausa intonation. In Kingston & Beckman (eds.), 17–34.
- Kingston, John & Mary E. Beckman (eds.). 1990. *Papers in Laboratory Phonology I: Between the grammar and the physics of Speech*. Cambridge: Cambridge University Press.
- Klatt, Dennis. 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3, 129–140.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2014. Tests in linear mixed effects models. <http://www.R-project.org>.
- Lenth, Russell, Henrik Singmann, Jonathon Love, Paul Buerkner & Maxime Herve. 2020. emmeans: Estimated marginal means, aka least-squares means. <https://cran.r-project.org/web/packages/emmeans/index.html>.
- Lieberman, Mark & Janet [B.] Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In Mark Aronoff & Richard Oehrle, with Frances Kelley & Bonnie Wilker Stephens (eds.), *Language sound structure*, 157–233. Cambridge, MA: MIT Press.
- Lindblom, Björn, Bertil Lyberg & Karin Holmgren. 1981. *Durational patterns of Swedish phonology: Do they reflect short-term memory processes?* Bloomington, IN: Indiana University Linguistics Club.
- McCarthy, John J. 1983. Consonantal morphology in the Chaha verb. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo García Pardo, Peter Guekguezian, Brian Hsu, Charlie O'Hara & Iris Chuoying Ouyang (eds.), *West Coast Conference on Formal Linguistics* 32 (WCCFL 32), 176–188. Somerville, MA: Cascadilla Proceedings Project.
- Mester, R. Armin & Junko Itô. 1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65, 258–293.
- Mücke, Doris & Anne Hermes. 2007. Phrase boundaries and peak alignment: An acoustic and articulatory study. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS XVI)*, Saarbrücken, 997–1000. <http://www.icphs2007.de>.
- Myers, Scott. 1996. Boundary tones and the phonetic implementation of tone in Chichewa. *Studies in African Linguistics* 25, 29–60.
- Myers, Scott. 1998. Surface underspecification of tone in Chichewa. *Phonology* 15, 367–391.
- Myers, Scott. 1999. Tone association and  $f_0$  timing in Chichewa. *Studies in African Linguistics* 28, 215–239.
- Myers, Scott. 2003.  $F_0$  timing in Kinyarwanda. *Phonetica* 60, 71–97.
- Myers, Scott. 2020. An acoustic study of sandhi vowel hiatus in Luganda. *Language and Speech* 63(3), 506–525.
- Myers, Scott & Benjamin B. Hansen. 2007. The origin of vowel length neutralization in final position: Evidence from Finnish speakers. *Natural Language & Linguistic Theory* 25, 157–193.
- Myers, Scott, Saudah Namyalo & Anatole Kiriggwajjo. 2019.  $F_0$  timing and tone contrasts in Luganda. *Phonetica* 76(1), 55–81.
- Myers, Scott, Elisabeth Selkirk & Yelena Fainleib. 2018. Phonetic implementation of high-tone spans in Luganda. *Laboratory Phonology* 9(1), 1–22.
- Pak, Marjorie. 2008. *The postsyntactic derivation and its phonological reflexes*. Ph.D. dissertation, University of Pennsylvania.
- Pierrehumbert, Janet B[reckenridge]. 1980. *The phonetics and phonology of English intonation*. Bloomington, IN: Indiana University Linguistics Club.
- Pierrehumbert, Janet B. & Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.
- Poser, William. 1984. *The phonetics and phonology of tone and intonation in Japanese*. Ph.D. dissertation, MIT.
- Prieto, Pilar, Jan van Santen & Julia Hirschberg. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23, 429–451.

- R Core Team. 2017. R: A language and environment for statistical computing (version 3.3.3). <http://www.R-project.org/>.
- Rialland, Annie. 2009. The African lax question prosody: Its realization and geographical distribution. *Lingua* 119, 928–949.
- Sadeghi, Vahid. 2017. The timing of pre-nuclear pitch accents in Persian. *Journal of the International Phonetic Association* 49(3), 305–329.
- Sibomana, Leonidas. 1974. *Deskriptive Tonologie des Kinyarwanda*. Hamburg: Helmut Buske.
- Silverman, Kim E. A. & Janet B. Pierrehumbert. 1990. The timing of prenuclear high accents in English. In Kingston & Beckman (eds.), 72–106.
- Smiljanić, Rajka. 2006. Early vs. late focus: Pitch-peak alignment in two dialects of Serbian and Croatian. In Louis Goldstein, Doug Whalen & Catherine Best (eds.), *Laboratory Phonology* 8, 495–518. Berlin: Mouton de Gruyter.
- Snoxall, R. A. 1967. *Luganda–English dictionary*. Oxford: Oxford University Press.
- Stevick, Earl. 1969a. Pitch and duration in Ganda. *Journal of African Languages* 8, 1–28.
- Stevick, Earl. 1969b. Tone in Bantu. *International Journal of American Linguistics* 35, 330–341.
- Thorsen, Nina. 1978. An acoustical investigation of Danish intonation. *Journal of Phonetics* 6, 151–175.
- Torreira, Francisco & Martine Grice. 2018. Melodic constructions in Spanish: Metrical structure determines the association properties of intonational tones. *Journal of the International Phonetic Association* 48, 9–32.
- Tucker, A. N. 1962. The syllable in Luganda: A prosodic approach. *Journal of African Languages* 1, 122–166.
- Turk, Alice. 1999. Structural influences on boundary-related lengthening in English. In John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville & Ashlee C. Bailey (eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences (ICPhS XIV)*, 237–240. Berkeley, CA: Linguistics Department, University of California at Berkeley.
- Yeou, Mohamed. 2004. Effects of focus, position and syllable structure on f0 alignment patterns in Arabic. Presented at Arabic Language Processing, Fez. [http://www.afcp-parole.org/doc/Archives\\_JEP/2004\\_XXVe\\_JEP\\_Fes/actes/arabe2004/PAMY11.pdf](http://www.afcp-parole.org/doc/Archives_JEP/2004_XXVe_JEP_Fes/actes/arabe2004/PAMY11.pdf)
- Ziervogel, Dirk, Petrus Wentzel & T. N. Makuya. 1972. *A handbook of the Venda language*, 2nd edn. Pretoria: UNISA.