

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Phonetics ■ (■■■■) ■■■-■■■

Journal of
Phoneticswww.elsevier.com/locate/phonetics

Downstep and high raising: interacting factors in Yoruba tone production

Yetunde O. Laniran^a, G.N. Clements^{b,*}^a1602 NE Junior Street, Portland, OR 97211, USA^bPhonetique Experimentale et Modelisation Phonologique, CNRS/Université de Paris 3, Sorbonne-Nouvelle ILPGA UMR 7018, 19 rue des Bernardins, Paris 75005, France

Received 25 January 2002; accepted 24 June 2002

Abstract

This study examines interacting factors in tone production in Yoruba, a tone language with three tone levels, high (H), mid (M), and low (L). Its primary goals are to confirm the existence of downstep, a principle which causes successive H tones separated by L tones to step down in pitch, and to examine the interaction between downstep and H tone raising, a principle which raises H tones to extra-high values before L tones. Controlled comparisons of data from four speakers reveal that both of these principles apply to H tones satisfying their conditions. As a result of H raising, the first H tone in downstepping sequences of the form HLHLH... is raised well above its expected value, while the following downstepped H tones are kept from descending into the frequency band reserved for M tones. This study also examines the strategies used for economizing pitch space in longer downstepping sequences. The main strategy used by all speakers is H tone resetting; however, some speakers are also found to raise initial H tones to extra-high values in anticipation of downsteps occurring four syllables away. Other interacting factors in Yoruba tone production include tone-specific declination (“downdrift”) operating in the background and local carry-over assimilation from a H tone to the following L tone. These various observations support a compositional model of tone production in which competing factors culminate on individual tones to produce functionally motivated “compromise” f0 patterns.

© 2003 Published by Elsevier Science Ltd.

1. Introduction

This study concerns the nature of tone production in Yoruba [yorùbá], a tone language spoken in Nigeria. The special interest of Yoruba is that it has three lexically distinctive tone levels, high

*Corresponding author. Tel.: +33-1-44-32-05-76; fax: +33-1-44-32-05-73.

E-mail addresses: ylaniran@aol.com (Y.O. Laniran), clements@idf.ext.jussieu.fr (G.N. Clements).

1 (H), mid (M) and low (L). Although three-level tone languages are not rare in Africa, to our
 2 knowledge, no other “register-tone system” (Pike, 1948) of this type has so far been studied in any
 3 detail with standard experimental techniques. As a result of this lacuna, studies of tone production
 4 in African tone languages continue to be heavily influenced by work on languages with two
 5 phonological tone levels, H and L. The experimental study of a three-level language such as
 6 Yoruba offers a potentially valuable new perspective on cross-linguistic principles underlying f0
 7 patterns.

8 The present paper addresses three closely related questions in tone production. The first
 9 involves the phonetic nature of downstep in languages, such as Yoruba, in which downstep is not
 10 lexically or grammatically distinctive. In tone languages, downstep typically consists of the
 11 stepwise lowering of H tones in linguistically defined contexts. In *automatic* downstep, H tones are
 12 lowered in sequences of alternating H and L tones; thus, for example, in a downstepping tone
 13 sequence HLHLH..., the second H tone is lower than the first, the third is lower than the second,
 14 and so forth. In *nonautomatic* downstep, often noted H¹H, there is no overt conditioning L tone
 15 between the two H tones.¹ Automatic downstep has been previously reported in Yoruba by
 16 Connell and Ladd (1990) and Laniran (1992). We attempt to confirm its existence and determine
 17 some of its properties in the first three experiments described below.

18 Downstep must be carefully distinguished from other principles that contribute to pitch
 19 lowering (see also Connell & Ladd, 1990, pp. 2–3). *Declination* is a continuous, long-term pattern
 20 of pitch decline across the utterance found commonly, though not exclusively, in nontone
 21 languages (Maeda, 1976; Pierrehumbert, 1980; Thorsen, 1983; Ladd, 1984). Unlike declination,
 22 downstep is localized at specific junctures and is usually conditioned by the tonal, lexical,
 23 morphological, and/or syntactic structure of the utterance in which it applies, often serving
 24 distinctive or demarcative functions. *Assimilation* is a term used by some phoneticians to refer to
 25 “the contextual variability of speech sounds, by which one or more of their phonetic properties
 26 are modified and become similar to those of the adjacent segments” (Farnetani, 1997, p. 376). By
 27 extension, this term has been used to describe the partial lowering of H tones in the context of
 28 lower tones; for example, Gandour, Potisuk, and Dechongkit (1994) use the terms “carry-over
 29 assimilation” and “carry-over coarticulation” interchangeably to describe the phonetic lowering
 30 of a tone with a high onset after a tone with a low offset. Downstep of H tones differs from
 31 assimilation in that it lowers the ceiling at which following H tones are realized; as a result, each
 32 successive H tone in longer downstepping sequences is lower than the preceding one, creating a
 33 cumulative “staircase” pattern. Downstep should also be distinguished from *final lowering*, an
 34 intonational effect which lowers the pitch of tones at the very end of (typically declarative)
 35 utterances (e.g. Liberman & Pierrehumbert, 1984; Pierrehumbert & Beckman, 1988). Downstep,
 36 unlike final lowering, applies most prominently at the beginning of the utterance.

37 A further pitch-lowering principle brought to light in Laniran’s earlier work on Yoruba (1992)
 38 is tone-specific declination, that is, declination applying to individual tone levels rather than to all
 39

40 ¹The widely-accepted terms *automatic* and *nonautomatic* downstep are due to Stewart (1965), who was the first to
 41 recognize an intimate relation between these two types of downtrend. While downstep normally affects H tones most
 42 prominently, L tones are distinctively downstepped in a few, rather atypical tone languages such as Kikuyu and
 43 Dschang-Bamileke, and it is a much-debated question whether L tones also undergo downstep in more representative
 tone languages as well. For earlier discussions of downstep in African languages, see Welmers (1973), Clements (1979),
 and Pulleyblank (1986), among others.

1 levels at once. Unlike downstep, tone-specific declination is not conditioned by other tones or by
 2 lexical or grammatical features. It is most clearly manifested in sequences of like tones. We will
 3 use the term “downdrift” in this paper as a convenient shorthand term for tone-specific
 4 declination, thus speaking of, e.g. “L tone downdrift” to refer to the declination of L tones
 5 independently of other tones.²

6 The concept of downstep has been extended to studies of many nontonal languages outside
 7 Africa, and has been incorporated in quantitative models of intonation in better-described
 8 languages such as English, Swedish, German, Dutch, French, and Japanese (see [Ladd, 1996](#);
 9 [Rialland, 1997](#) for reviews). It is perhaps surprising that relatively less quantitative work has been
 10 carried out on the African tone languages from which the concept of downstep originally
 11 emanated. Phonetic studies of African tone languages other than Yoruba include those of
 12 [Mountford \(1983\)](#) on Bambara and of [Lindau \(1986\)](#) and [Inkelas and Leben \(1990\)](#) on Hausa. An
 13 advantage of studying downstep in African tone languages is that surface-phonological tone
 14 sequences and their patterns of alignment to vowels are relatively well understood and allow tone-
 15 f₀ correspondences to be studied with some precision.

16 A related concern of this study is to determine how downstep interacts with other pitch-scaling
 17 principles in Yoruba, particularly H tone raising, according to which H tones are realized with
 18 extra-high f₀ values before L tones. As [Connell and Ladd \(1990, pp. 18–19\)](#) have observed, the
 19 apparent downstepping of the second H tone in HLH sequences might in some cases be analyzed
 20 instead as H raising of the first, which would give an effect similar to downstep; as they point out,
 21 their Yoruba corpus did not provide a systematic basis for distinguishing the two phenomena. H
 22 raising seems to have been first reported in African languages by [Pike \(1966\)](#), and has
 23 subsequently been confirmed experimentally in several tone languages of Africa and Asia
 24 including Gurma ([Rialland, 1981](#)), Yoruba ([Connell & Ladd, 1990](#); [Laniran, 1992](#); [Akinlabi &](#)
 25 [Liberman, 1995](#)), Thai ([Gandour et al., 1994](#)), Igbo ([Laniran & Gerfen, 1997](#)), Mandarin Chinese
 26 ([Xu, 1997](#)), and Bimoba ([Snider, 1998](#)). The phonologization of raised H tones often gives rise to
 27 new surface-contrastive tone levels ([Rialland, 1983](#); [Hyman, 1993](#)). However, its interaction with
 28 downstep and other pitch-scaling principles is still poorly understood.

29 A further goal of this study is to determine whether downstep production involves “foresight”
 30 or “hindsight” in Yoruba. Given that any speaker’s pitch range is not infinitely expandable, we
 31 expect that speakers will adopt some strategy to economize pitch space in longer downstepping
 32 utterances. Strategies employing foresight include initializing the first H in the sentence at an
 33 extra-high f₀ value to make space for following downsteps ([Rialland, 2001](#)). Hindsight strategies,
 34 in contrast, are brought into play only when speakers run out of pitch space; for example,
 35 speakers sometimes reset their H tone values upward after they have fallen to a certain minimum
 36 (see [Berg, Gussenhoven, & Rietveld, 1992](#)).

37 The issue of whether speakers of tone languages ever employ “foresight” in producing longer
 38 downstepping sequences has been a matter of controversy. [Stewart \(1965\)](#) has claimed that the
 39 pitch of H tones in downstepping sequences in Akan is sensitive to the number of following

41 ²The term “downdrift” has sometimes been used to refer indifferently to automatic downstep or declination, which
 42 have not always been consistently distinguished in the earlier literature on tone languages. The use of this imprecise
 43 term with its confusing range of uses has been severely and rightly criticized by [Connell and Ladd \(1990, pp. 2–3\)](#). We
 use it here only in the specific sense described above.

1 downsteps. He states that “the pitch of any particular high tone is raised by as many levels as
 2 there are downsteps in the subsequent part of the phrase,” while the last H tone in the sequence
 3 tends to be realized at a constant level, its “basic pitch.” Contesting this claim, also on the basis of
 4 impressionistic data, Schachter maintained that the pitch of the first H tone in Akan is “normally
 5 phonetically the same” regardless of the number of following downsteps, while later H tones
 6 descend to lower and lower values as the number of downsteps increases (Schachter, 1965).
 7 Schachter’s view rejects look-ahead as a feature of Akan tone production, and has represented the
 8 dominant trend in subsequent descriptive work on most other African downstep systems.
 9 However, only recently have experimental studies been conducted to test for possible “look-
 10 ahead” effects.³ In preliminary reports on the experiment described in Section 6 of this paper,
 11 Laniran (1992) and Laniran and Clements (1995) have reported look-ahead effects in Yoruba
 12 tone production, though not of the magnitude and consistency of those which Stewart reported.
 13 More recently, Rialland and Somé (2000) have described long-distance look-ahead effects in
 14 Dagara, some of which resemble the pattern noted by Stewart quite closely.

15 The remainder of this article is organized as follows. Section 2 provides background
 16 information on Yoruba tone structure and outlines our methodology. Sections 3–6 are structured
 17 around a series of experiments designed to shed light on the questions raised above. Experiment 1,
 18 described in Section 3, brings to light a pattern of apparent downsteps affecting both H and L
 19 tones in “mixed-HL” sentences, and shows that speakers typically use resetting (a hindsight
 20 strategy) to economize pitch space in longer sentences. Experiment 2 (Section 4), using controlled
 21 comparisons, confirms that the downtrends observed in Experiment 1 data cannot be attributed to
 22 independent utterance level or tone-specific declination. It also shows that H tones have
 23 unexpectedly high values in mixed-HL sentences when compared with the same tones in sentences
 24 containing only H tones, an effect which is attributed to H raising. Experiment 3 (Section 5)
 25 examines and rejects an alternative account of this raising effect, according to which H tones are
 26 not raised in mixed-HL sentences but globally lowered in all-H sentences. This section also shows
 27 that M tones are neither triggers nor targets of downstep. Experiment 4 (Section 6) reports a long-
 28 distance “foresight” effect in sentences beginning with the sequence HLHY (Y = H, M, or L),
 29 according to which the first H has a significantly higher than expected value for some speakers if
 30 tone Y is an L tone. Section 7 discusses two alternative interpretations of Experiments 1–4 data
 31 which do not require positing an interaction between downstep and H raising, and shows that
 32 neither of them accounts for certain crucial facts. Section 8 discusses quantitative aspects of
 33 downstep, H raising, and resetting, and shows that downstep implementation strategies vary
 34 considerably according to speaker and context. Section 9 concludes that the Yoruba data
 35 motivate a compositional model of tone production in which several interacting factors culminate
 36 on individual tones and tone levels to produce the observed f₀ patterns.
 37
 38
 39
 40
 41

42
 43 ³To our knowledge, the dispute between Stewart and Schachter has not yet been resolved by instrumental studies of
 Akan.

2. Background information

The tonal phonology of Yoruba is relatively well known through the descriptions of Ward (1952), Akinlabi (1985), and others. However, while different aspects of tonal phonetics have been examined by Hombert (1977), Connell and Ladd (1990), Laniran (1992, 1995), Laniran and Clements (1995), Akinlabi and Liberman (1995), and others, many basic questions remain unanswered. For example, though earlier studies have brought to light the presence of both downstep and H raising in Yoruba (Connell & Ladd, 1990; Laniran, 1992), the interaction between these two phenomena, if they are indeed independent, remains unclear. One task, then, will be to sort out the relative contribution of downstep and H raising to the overall pattern of f₀ decline across H tones, and to determine whether both of these factors are independently needed to account for phonetic generalizations. Another unresolved question concerns whether speakers use “foresight” or “hindsight” strategies, as discussed in the introduction, to accommodate large numbers of downsteps in longer utterances. Further questions, specific to downstep in languages like Yoruba with three (or more) tone levels, include: Does downstep involve just H tones, or does it affect M and L tones as well? Does downstep affect H tones separated only by L tones, or may M tones intervene?

This section reviews the basic aspects of Yoruba tone structure and then presents our experimental method.

2.1. Yoruba tonal structure

Yoruba is a tone language with three lexically distinctive tone levels, as illustrated in (1). In our transcriptions, H (high) tones are marked with the acute accent (´), L (low) tones are marked with the grave accent (`), and M (mid) tones are unmarked:

lú	‘to mix together,’	ilú	‘city,’	
lu	‘to perforate,’	ilu	‘perforator,’	(1)
lù	‘to beat,’	ilù	‘drum.’	

A further aspect of Yoruba tonology essential to the understanding of tone production is Tone Spread. Whenever a H tone syllable and a L tone syllable come into contact (in either order), the tone of the first syllable spreads onto the second to create a contour tone. For example, the proper name *Máyò mí*, bearing the lexical tone sequence H–L–H, is regularly realized with a falling (HL) tone on its second syllable and a rising (LH) tone on its third. The effect of Tone Spread is indicated by the dashed lines in the following representation of the sentence *Máyò mí ra ìwé* ‘Máyò mí bought books’:⁴

Má	yò	mí	rà	wé	= [má yò mǐ rà wě]	
						(2)
└─┘	└─┘	└─┘	└─┘	└─┘		
H	L	H	L	H		

Each tone in this sequence lags onto the following vowel, turning lexical L tones into falling tones

⁴In normal pronunciation, the elided initial vowel *i* of *ìwé* passes its H tone, the preceding vowel (an effect known as “tone stability”).

1 from that indicated in the orthography were discarded. The remaining recordings were digitized at
 2 a sampling rate of 10 kHz using the Entropics Waves+ software on a SUN Workstation (at
 3 Cornell) or (in the case of FA) the GWI Soundscope software for the Macintosh (at the University
 4 of North Carolina at Chapel Hill). After elimination of sentences with incomplete f0 tracks, ten
 5 tokens of each sentence from TJ and BJ and nine tokens by KG and FA were retained for
 6 analysis. F0 was extracted using the auto-correlation technique (Rabiner & Schafer, 1975).

7 Segmentation of the signal was carried out by visual comparison of the f0 tracings with
 8 synchronized spectrograms and waveforms, and by playback of selected segments for auditory
 9 verification. In order to take into account the effect of Tone Spread which, as noted in the
 10 preceding section, creates sharply falling f0 contours on H and L syllables, two measurement
 11 points were selected in each vowel. The first point (p_1) was selected toward the beginning of the
 12 vowel, and the second (p_2) within the final 25 ms. However, measurements were taken only in
 13 regions not showing any notable consonant-induced perturbations of the f0 contour as discussed
 14 above.

17 3. Experiment I

19 The first experiment was designed to find out whether Yoruba sentences display downstep
 20 effects across H and L tones. The following questions were addressed:

- 21 1. Do H tone sequences show downstep effects?
- 22 2. Do L tone sequences show downstep effects?
- 23 3. Does a sentence-initial L tone downstep a following H tone?

25 The test sentences used in this experiment also provide data bearing on the issue of
 26 preplanning. As discussed earlier, if speakers preplan the f0 structure of longer utterances in
 27 order to economize pitch space, we might expect that as the number of downsteps in the
 28 downstep span increases, initial H tones will shift upward, or the downstep interval will
 29 decrease. Thus, the following further questions were addressed:

- 30 4. Are initial H tones scaled higher as the number of downsteps increases?
- 31 5. Does the size of the downstep interval decrease as the number of downsteps increases?

32 If, on the other hand, there is no preplanning, we might expect that as the number of
 33 downsteps in the downstep span increases, final H values will drop to lower and lower values,
 34 or that some later H tones will be reset to higher values:

- 35 6. Do final H tones drop to lower values in sentences with more downsteps?
- 36 7. Are later H tones reset to higher values in sentences with more downsteps?

37 3.1. Corpus design

39 In order to address these questions, we designed two sets of sentences in which H and L tones
 40 alternate on successive syllables (see Appendix A). We call these “mixed-HL sentences.” Each set
 41 contained eight sentences, varying in length. One set of sentences, which we call “L-initial
 42 sentences,” started with a L tone; the other set, containing “H-initial sentences,” started with a H
 43 tone. The two sets were tonally identical in all other respects; thus the tonal pattern of the L-initial

1 sentences differs from that of H-initial sentences of the same length only in having an extra L tone
 2 syllable at the beginning.

3

5 3.2. Results

7 **Fig. 1** graphs the mean f_0 values of the first four L-initial sentences for speaker TJ.

8 In these graphs, the x -axis displays syllables labelled by the lexical tones they bear, and the y -
 9 axis displays fundamental frequency (f_0) in Hz. As described earlier, two measurements are given
 10 for each syllable. For example, **Fig. 1a** shows the first (shortest) sentence. The first two f_0 values in
 11 this graph are those obtained for points p_1 and p_2 in the first syllable (bearing a lexical L tone), the
 12 next two values are those for points p_1 and p_2 in the second syllable (bearing a lexical H tone), and
 13 so on. Due to the effect of Tone Spread (2), each lexical tone in these sentences is realized not only
 14 on the syllable with which it is directly aligned but also at the beginning of the following syllable
 15 as well. Thus, for example, the initial L tone is represented by the first three points on the graph,
 the following H tone by the next two, and so on.⁵

16 In the interest of clarity, such graphs can be simplified by selecting only the lowest value of each
 17 L tone and the highest value of each H tone, whether these values are aligned with the syllable of
 18 origin or with the following syllable as a result of Tone Shift. In this way, peaks and valleys can be
 19 displayed in abstraction from their alignment with particular syllables. **Figs. 2–4** show patterns of
 H and L extrema obtained by this method.

20 These figures superimpose graphs of averaged L-initial and H-initial test sentences as spoken by
 21 TJ (**Fig. 2**), KG (**Fig. 3**), and FA (**Fig. 4**), respectively. Only H tones are labelled on the x -axis; the
 22 intervening values represent L tones.

23 Let us take up the questions raised at the beginning of this section.

25

- 26 1. *Do H tone sequences show downstep effects?* The data reveal a prominent cumulative lowering of
 27 H tones strongly resembling downstep throughout the short sentences (a–c) and at the
 beginning of longer sentences (d–h). This effect extends over the first three H tone peaks for
 28 KG, the first three or four for FA, and over as many as six for TJ.
- 29 2. *Do L tone sequences show downstep effects?* Speakers TJ and KG, but not FA, show some
 30 lowering across the second and third, and sometimes fourth, L tones, especially in short
 31 sentences and near the beginnings of longer ones. However, later L tones show no consistent
 32 lowering pattern, even after “reset” H tones (see below). These two speakers reveal a further,
 33 unexpected effect: the first L tone in each of their L-initial sentences is much lower than the
 34 following one, and may be as low as the third or fourth L tones.
- 35 3. *Does a sentence-initial L tone downstep a following H tone?* Here again we find variation among
 36 our speakers. For TJ and KG (**Figs. 2 and 3**), the initial L tone in the L-initial sentences has no

37
 38
 39 ⁵Two irregularities in these graphs require comment. First, both here and in **Figs. 2–4**, the third H tone of graph (b) is
 40 lower than expected, due to the influence of a floating M tone in this test sentence as discussed in Appendix A. Second,
 41 the fourth L tone in graph (d) fails to spread as expected to the first measurement point in the following H tone syllable.
 Examination of the test sentence in question shows that this H tone occurs on a vowel sequence bearing a HL contour
 42 tone (*fún Ì*, phonetically [fúì]). We suggest that the H tone may be prevented from spreading onto the H tone syllable in
 43 this case by a “no-crowding” principle which prohibits vowel sequences from bearing more than two phonological
 tones.

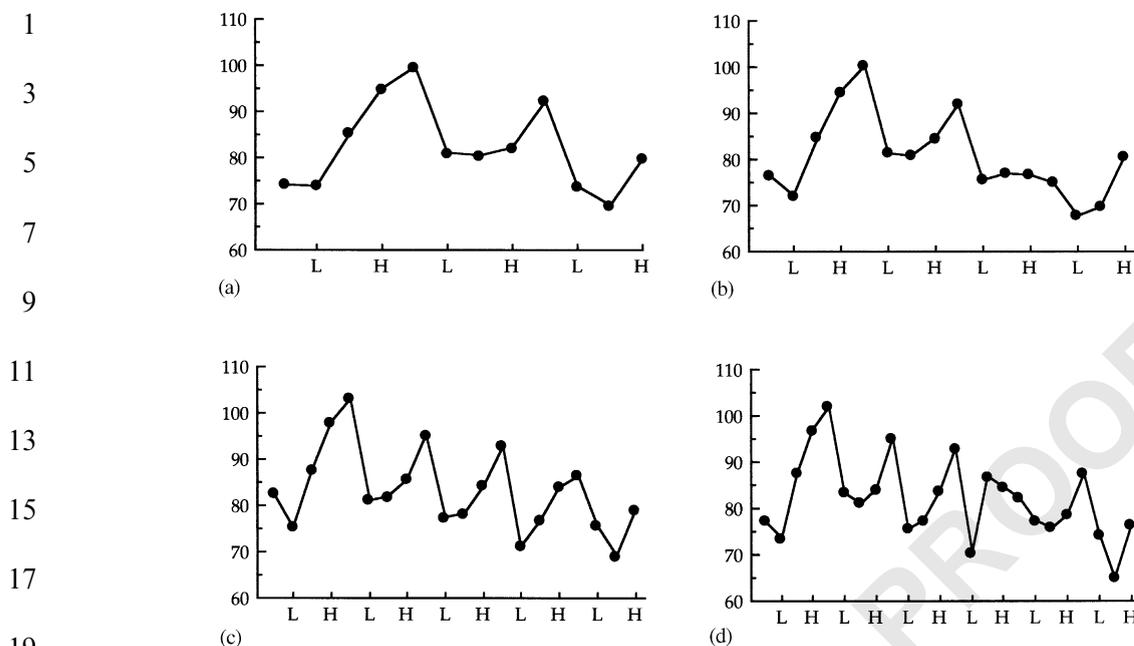


Fig. 1. Downstepping patterns in Yoruba. Mean f_0 values in Hz of four mixed-HL sentences, differing in length, are shown for speaker TJ. Each syllable, labelled along the x -axis by its lexical tone, is represented by two data points on the graph, one directly over it and the other just to its left. These patterns reflect the operation of Tone Spread, which spreads each lexical H or L tone onto the beginning of the following syllable.

lowering effect on the following H tone values. Indeed, for KG, it tends to have a small raising effect; a t -test showed this result to be just significant at the 0.05 level ($p < 0.0438$). For FA (Fig. 4), the opposite is true: the first few H tones in L-initial sentences are significantly lower than those in H-initial sentences ($p < 0.0194$). However, this difference is not of the magnitude we would expect from downstep; if genuine downstep were at work in FA's data, we would expect the first H tone in her L-initial sentences to have roughly the same value as the second H tone in her H-initial sentences. We conclude, with Connell and Ladd (1990), that sentence-initial L tones do not trigger downstep for any of our speakers. The slight lowering of the first few H tones in FA's L-initial sentences must be attributed to some other factor.

The next questions bear on the issue of preplanning.

4. Are initial H tones scaled higher as the number of downsteps increases? Visual inspection of Figs. 2–4 reveals no trend in this direction. Table 1 gives average values in Hz of the first H tone (H_1), the second H tone (H_2), and the final H tone (H_f), and the pitch drop between first and second H tones ($H_1 - H_2$) in the L- and H-initial sentences for speakers TJ, KG, and FA.

If initial H tones were scaled higher as the number of downsteps increases, the values of H_1 should rise steadily as we scan the H_1 row from left to right. It can easily be seen that this is not the case. It will be noted from Figs. 2–4, however, that many of the later H tones in the longer sentences (d)–(h) are reset to higher values, initiating new *downstep spans* (maximal sequences containing only descending H tone values). Preplanning is at issue, of course, only in single downstep spans. To test the possible preplanning effects on H_1 , therefore, we must count

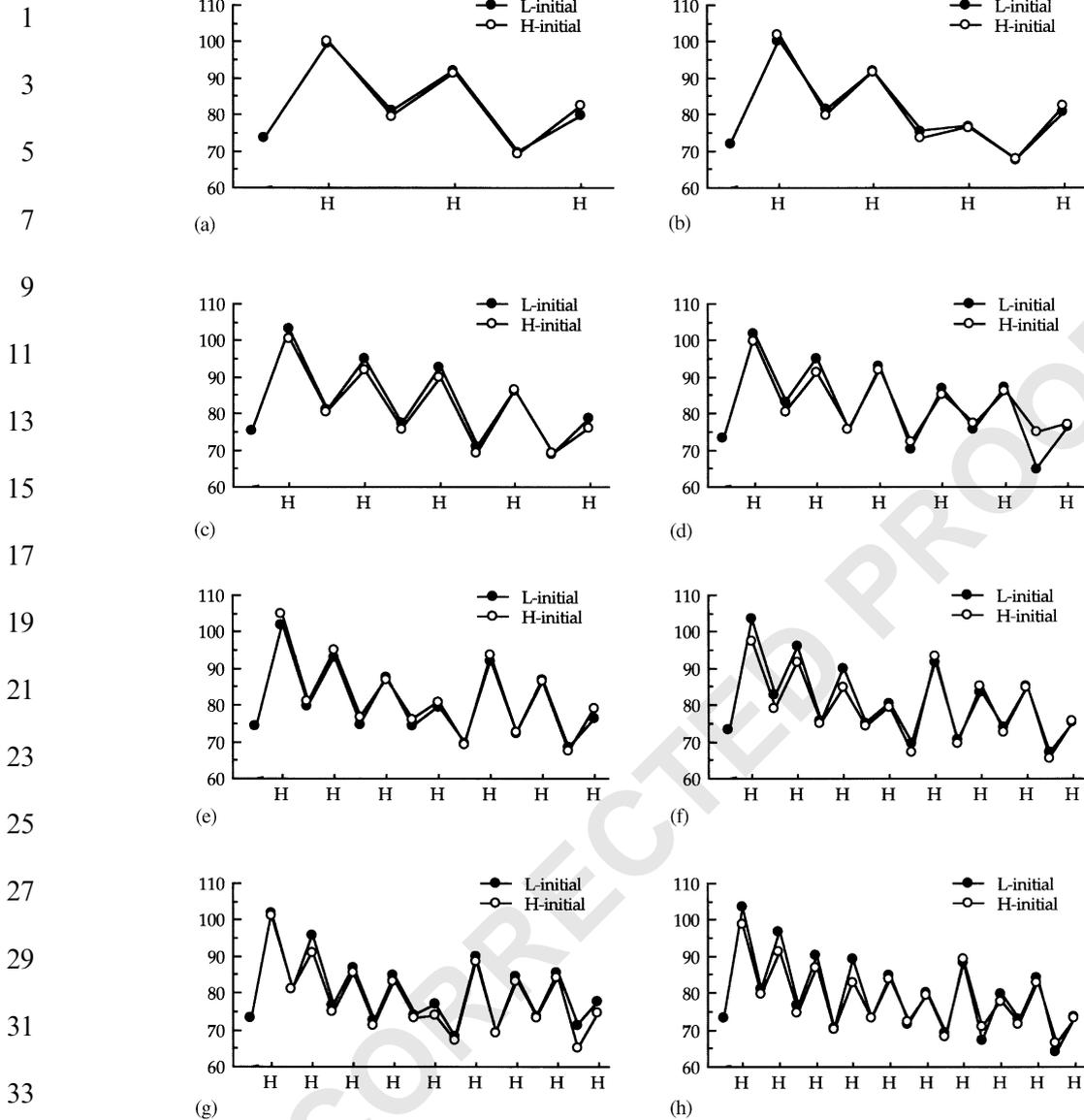


Fig. 2. Downstep patterns of mixed-HL sentences (speaker TJ). Values for L-initial and H-initial sentences are superimposed (values in Hz).

downsteps only in the initial downstep span. Regression analysis performed on the full data set shows a significant correlation only in TJ's L-initial sentences, tracked by black circles in Fig. 2, where the value of the initial H tone is set about 1 Hz higher for each additional downstep in the first downstep span ($y = 0.97x + 99.19$, $r^2 = 0.641$, $p = 0.017$). A more thorough study of "look-ahead" in this speaker's tone production will be undertaken in the presentation of Experiment 4 (Section 6). Other speakers showed no significant trend in this direction.

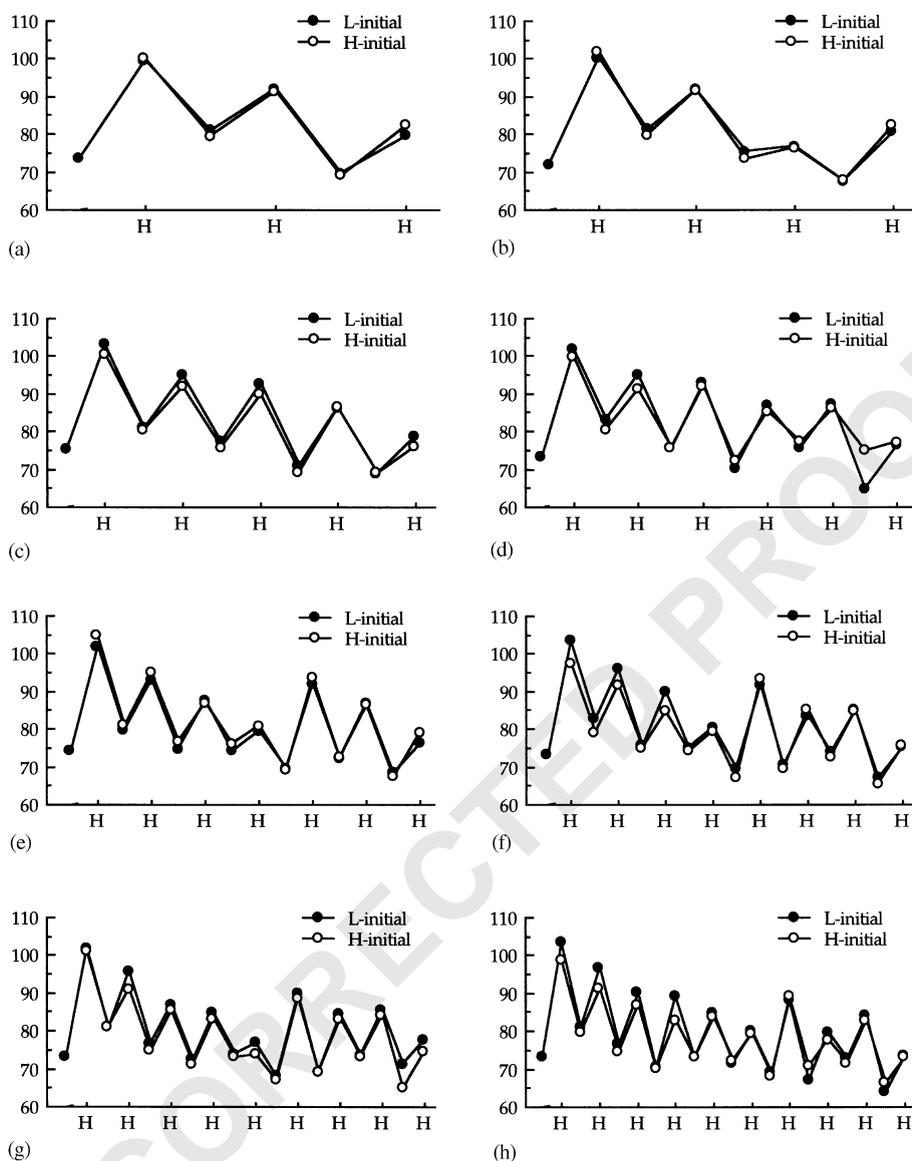


Fig. 3. Downstep patterns of mixed-HL sentences (speaker KG). Values for L-initial and H-initial sentences are superimposed (values in Hz).

5. Does the size of the downstep interval decrease as the number of downsteps increases? Again, Table 1 shows that this is not the case; scanning the values of the lowering interval $H_1 - H_2$ from left to right across Table 1, we see that they do not steadily decrease as sentences become longer, for any speaker. Regression analysis confirms the lack of any significant effect.
6. Do final H tones drop to lower values in sentences with more downsteps? Table 1 shows an effect of sentence length on final H tone (H_f) values only for speaker TJ, and a minimal one at that. Regression analysis shows that the values of TJ's final H tones tend to drop by roughly 1 Hz

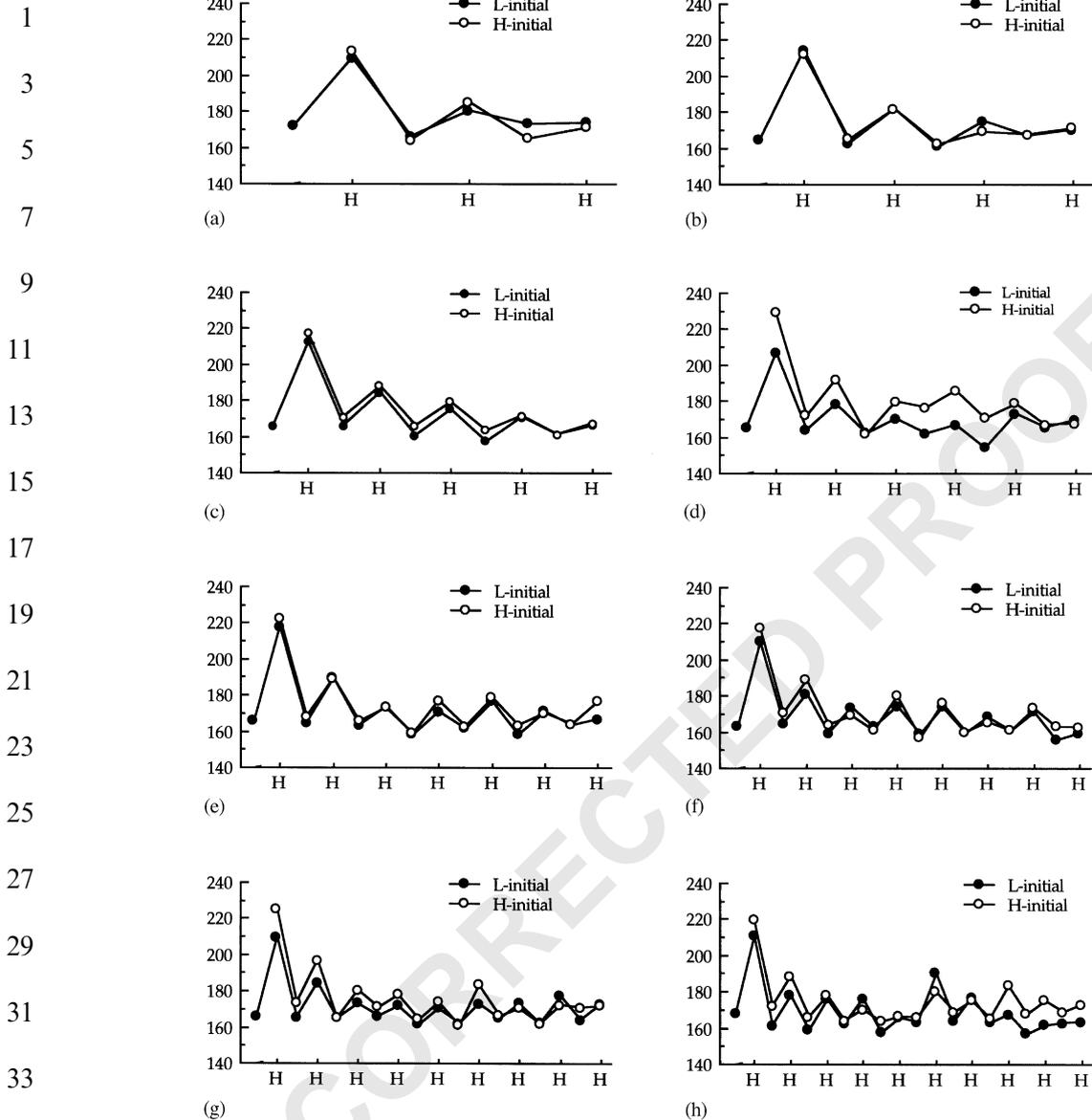


Fig. 4. Downstep patterns of mixed-HL sentences (speaker FA). Values for L-initial and H-initial sentences are superimposed (values in Hz).

with each increase in length, from a value of about 81 Hz in the shortest sentences to about 74 Hz in the longest ones ($y = -1.013x + 84.198$, $r^2 = 0.821$, $p = 0.0019$). The correlation remains significant if we use the number of effectively downstepped H tones (thus excluding any reset H tones) as the independent variable ($y = -1.672x + 82.18$, $r^2 = 0.431$, $p = 0.0057$). For the other speakers, there is no significant correlation between sentence length and H_f . KG's

1 Table 1

3 Average values of first, second, and final H tones in all L- and H-initial mixed-HL sentences for three speakers, with
 5 means and standard deviations for each

	Sentence	a	b	c	d	e	f	g	h	Mean	S.D.
	<i>n</i> (H)	3	4	5	6	7	8	9	10		
	<i>TJ (L-initial)</i>										
7	H_1	99.7	100.4	103.2	102.0	102.1	103.6	102.1	103.7	102.1	1.4482
	H_2	92.3	92.1	95.2	95.3	93.2	96.1	95.9	97.0	94.6	1.8531
9	H_1-H_2	7.4	8.3	8.0	6.7	8.9	7.5	6.2	6.7	7.5	0.8529
	H_f	79.8	80.8	79.0	76.4	76.4	75.4	78.0	73.7	77.4	2.3870
11	<i>TJ (H-initial)</i>										
	H_1	100.3	102.1	100.8	99.8	104.9	97.6	101.2	98.8	100.7	2.2055
13	H_2	91.6	91.9	92.1	91.3	95.1	91.9	91.2	91.4	92.1	1.2682
	H_1-H_2	8.7	10.2	8.7	8.5	9.8	5.7	10.0	7.4	8.6	1.4069
15	H_f	82.5	82.6	76.3	77.2	79.3	75.7	74.9	73.3	77.7	3.4437
	<i>KG (L-initial)</i>										
17	H_1	148.5	149.9	148.5	152.2	151.6	147.5	149.6	147.4	149.4	1.7719
	H_2	120.9	128.4	128.7	129.6	126.2	125.9	128.5	128.9	127.1	2.8380
19	H_1-H_2	27.6	21.5	19.8	22.6	25.4	21.6	21.1	18.5	22.3	2.7622
	H_f	83.5	89.4	83.4	82.9	81.7	86.3	83.0	83.0	84.2	2.4894
21	<i>KG (H-initial)</i>										
	H_1	145.3	145.6	150.6	147.2	144.5	148.8	145.9	145.6	146.7	2.0608
23	H_2	121.1	121.6	125.6	128.1	126.6	125.8	119.1	121.8	123.7	3.2011
	H_1-H_2	24.2	24.0	25.0	19.1	17.9	23.0	26.8	23.8	23.0	2.8003
25	H_f	81.4	86.9	78.7	82.2	80.5	83.4	79.4	81.3	81.7	2.5678
	<i>FA (L-initial)</i>										
27	H_1	210.0	214.8	213.1	207.1	217.9	210.2	209.9	210.8	211.7	3.3765
	H_2	180.7	181.9	184.8	178.6	190.0	181.3	184.8	178.4	182.6	3.8404
29	H_1-H_2	29.3	32.9	28.3	28.5	27.9	28.9	25.1	32.4	29.2	2.3404
	H_f	174.1	170.4	166.8	169.7	167.2	159.6	173.1	163.6	168.1	4.8397
31	<i>FA (H-initial)</i>										
	H_1	213.8	212.4	217.7	229.4	222.9	217.8	225.2	219.6	219.9	5.7629
33	H_2	185.4	182.0	188.3	192.2	189.4	189.0	196.9	188.7	189.0	4.4024
	H_1-H_2	28.4	30.4	29.4	37.2	33.5	28.8	28.3	30.9	30.9	2.8801
35	H_f	171.5	171.6	167.8	167.8	176.9	163.1	172.3	173.0	170.5	4.1906

37 Standard deviations were computed over individual tokens. *n*(H)=number of H tones.

39

41

43

final H tones vary around a mean of 83.0 Hz (S.D. = 2.30), and FA's around a mean of 169.3 Hz (S.D. = 3.57).

7. Are later H tones reset to higher values in sentences with more downsteps? Here the answer is affirmative: resetting is the major strategy used by all speakers to economize pitch space. All speakers reset the H tone "ceiling" upward at least once in the longer sentences, and do so as

1 many as three times in the longest sentence (h). After each resetting, the downstep pattern
 2 generally resumes over the next one or two H tones, though on a reduced scale. (A quantitative
 3 evaluation of these data is presented in Section 8.3.)

5

7

9

11

13

15

17

19

21

23

25

4. Experiment 2

27

29

31

33

35

37

39

41

43

The next experiment addresses the question whether the lowering patterns brought to light in Experiment 1 represent genuine downstep, or some other effect. A declining pitch ramp across H tones could result from other principles such as general background declination of all tones, or downdrift of H tones alone. In order to eliminate these alternative explanations, we must confirm that the lowering pattern is indeed triggered by the presence of intervening L tones. To do this, we may compare f_0 values of H tones in mixed-HL sentences with those of comparable H tones in sentences which lack L tones (cf. Beckman & Pierrehumbert, 1992, p. 391). If H tones show a similarly declining ramp in such sentences, we would be forced to conclude that some principle other than downstep is at work in both data sets.

Accordingly, we conducted a further experiment with speakers TJ and FA to see how the H tones in mixed-HL sentences compare with H tones in sentences of similar length and structure which lack the intervening L tones. To do this, we constructed two new sets of sentences, one containing only H tones (“all-H sentences”) and the other only L tones (“all-L sentences”). Earlier studies of Yoruba have reported that all-H sentences exhibit no appreciable pitch decline and that all-L sentences show a small pitch decline (Hombert, 1974; Connell & Ladd, 1990; Laniran, 1992). Therefore, if downstep is really the principle at work in the mixed-HL sentences of Experiment 1, we expect H tones (and perhaps L tones as well) to decline at a faster rate in these sentences than in the all-H (and all-L) sentences.

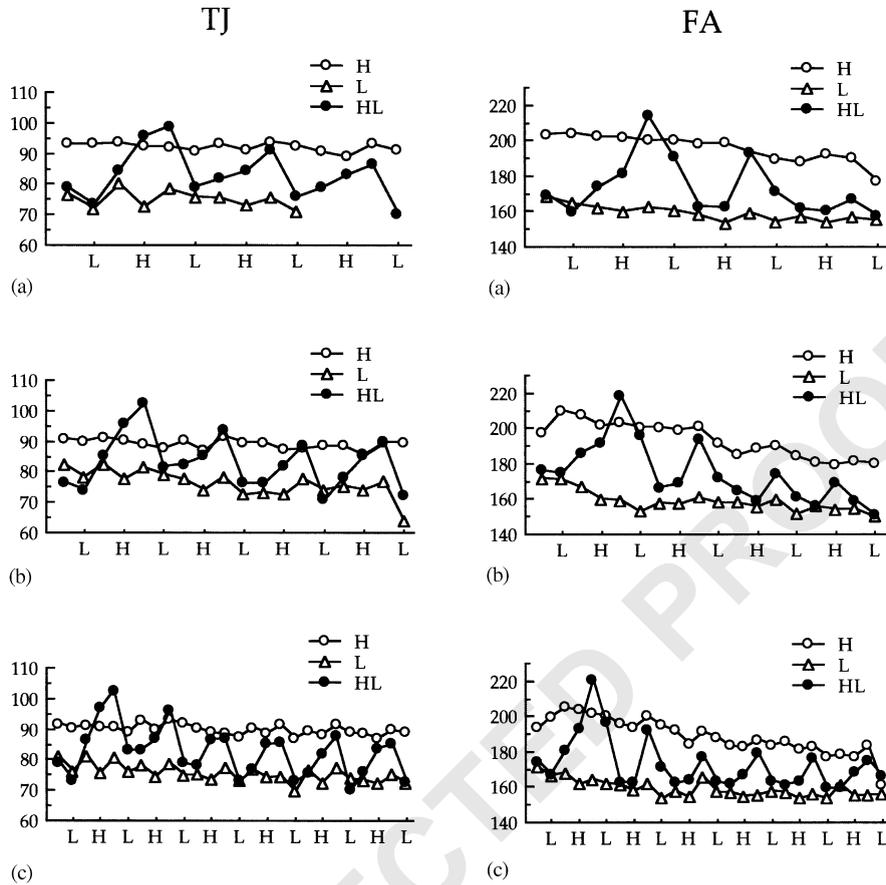


Fig. 5. Superimposed graphs of all-H, all-L and mixed-HL sentences superimposed (left column: TJ, right column: FA), showing that L tones downstep H tones (values in Hz).

4.1. Corpus design

The test sentences used in this experiment fell into three sets: all-H sentences, all-L sentences, and mixed-HL sentences similar to those used in Experiment 1. These sentences had a syntactic structure similar to those used in the earlier experiment. There were three length conditions in each set: (a) long, (b) medium, and (c) short. The test sequences were placed at the beginning of carrier sentences to absorb any final lowering effects. The sentences used in this experiment are listed in Appendix B. All sentences were read by TJ and FA and recorded and analyzed as described earlier.

4.2. Results

Fig. 5 overlays the graphs of mixed-HL sentences with those of the all-H and all-L sentences (tones of the sentence-final carrier frames are omitted here and in later figures). Graphs for TJ are shown on the left and those for FA on the right. Sentences are presented from shortest (a) to

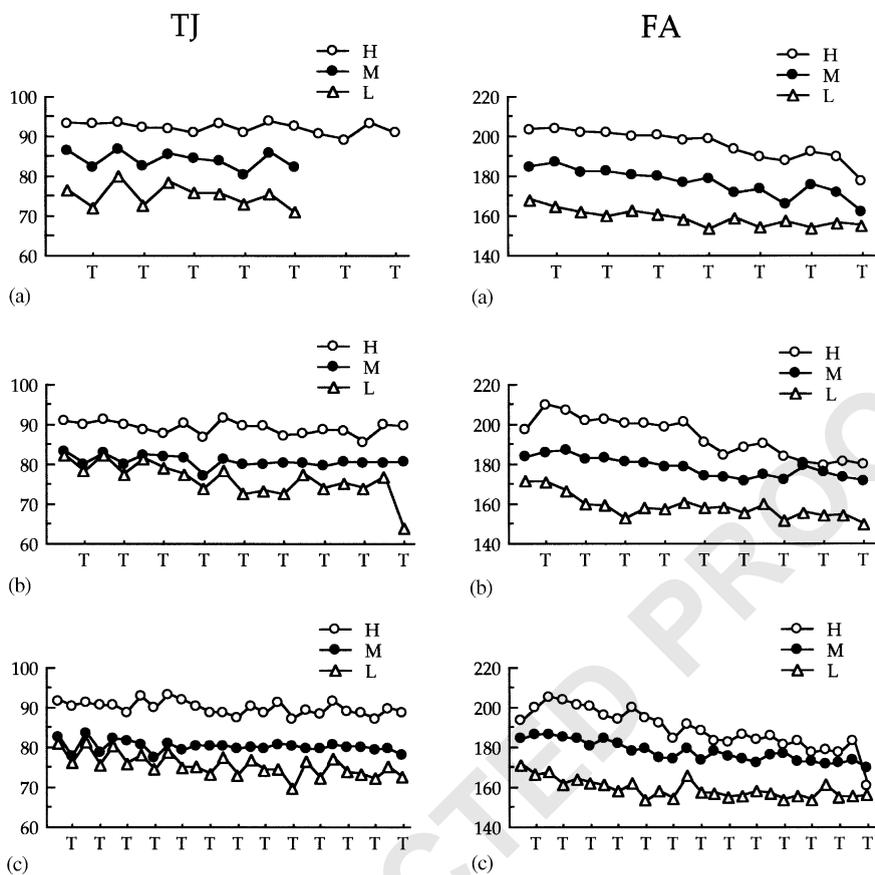


Fig. 6. Superimposed graphs of all-H, all-M, and all-L sentences for TJ and FA, showing differing rates of declination across different tones (values in Hz).

longest (c).⁶ Each graph shows average values for all tokens. Two measurement points (p_1 and p_2 as discussed in Section 2.2) are shown for each syllable, aligned as in Fig. 1. Syllables are identified by tone symbols corresponding to the H and L tones of the mixed-HL sentences. These symbols should be replaced across the board by H and L tone symbols as appropriate in interpreting the all-H and all-L sentences.

It can be seen that in all cases, H tones descend more rapidly in the mixed-HL sentences (tracked with black circles) than in the all-H sentences (tracked with white circles). Considering the patterns in more detail, we see that the first H tone of the mixed-HL sentences is considerably higher, in all graphs, than the first H tone of the all-H sentences. In the longer sentences (b) and (c), TJ's second H peak is also somewhat higher in the mixed-HL sentences. In several of the graphs, later H tones in the mixed-HL sentences descend to values well under those of the

⁶Due to an oversight in corpus design, the shortest all-L and all-M sentences (a) read by TJ contained only five syllables, explaining the absence of values here and in Figs. 6–8 for the last two syllables of the (a) graphs for this speaker. This oversight was corrected in FA's test sentences.

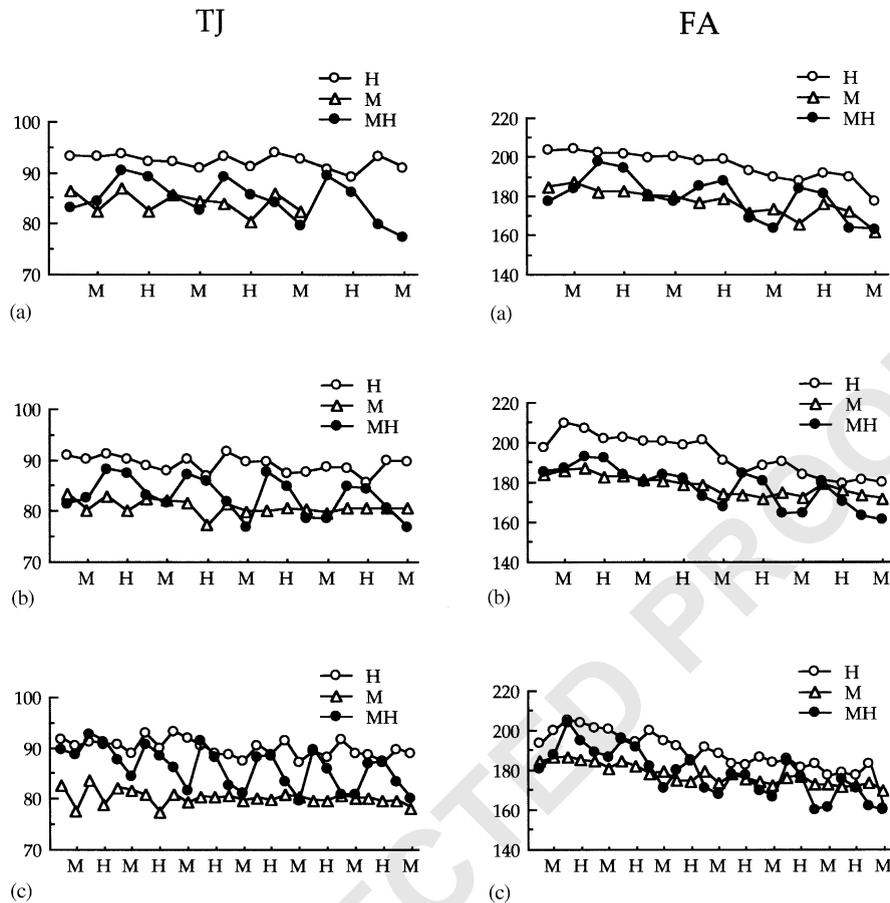


Fig. 7. Superimposed graphs of mixed-MH, all-M and all-H sentences for TJ and FA, showing that M tones do not downstep H tones (values in Hz).

comparable tones in the all-H sentences. These comparisons make it apparent that the sharply descending H tones in mixed-HL sentences cannot be accounted for in terms of some kind of general H-tone declination, but crucially depend on the presence of intervening L tones, the defining condition of downstep.

Comparison of the L-tone patterns leads to a different conclusion. We observe a general declining pattern across L tones in all of TJ's mixed-HL sentences, as well as in FA's mid-length sentences (b). However, the L tones in the mixed-HL sentences do not descend below the values of the comparable L tones in all-L sentences (labeled with white triangles). In fact, they often remain well above them, except for TJ's sentence-initial L's which, as in his mixed-HL sentences of Experiment 1, start at an unusually low value. We thus have no evidence that L tones undergo downstep in mixed-HL sentences; the L-tone patterns in these sentences are very similar to those of the all-L sentences. If downstep means a drop below comparable nondownstepped values, most of the L tones cannot be said to be downstepped in these data. (We return to a discussion of L-tone patterns in Section 5.3.)

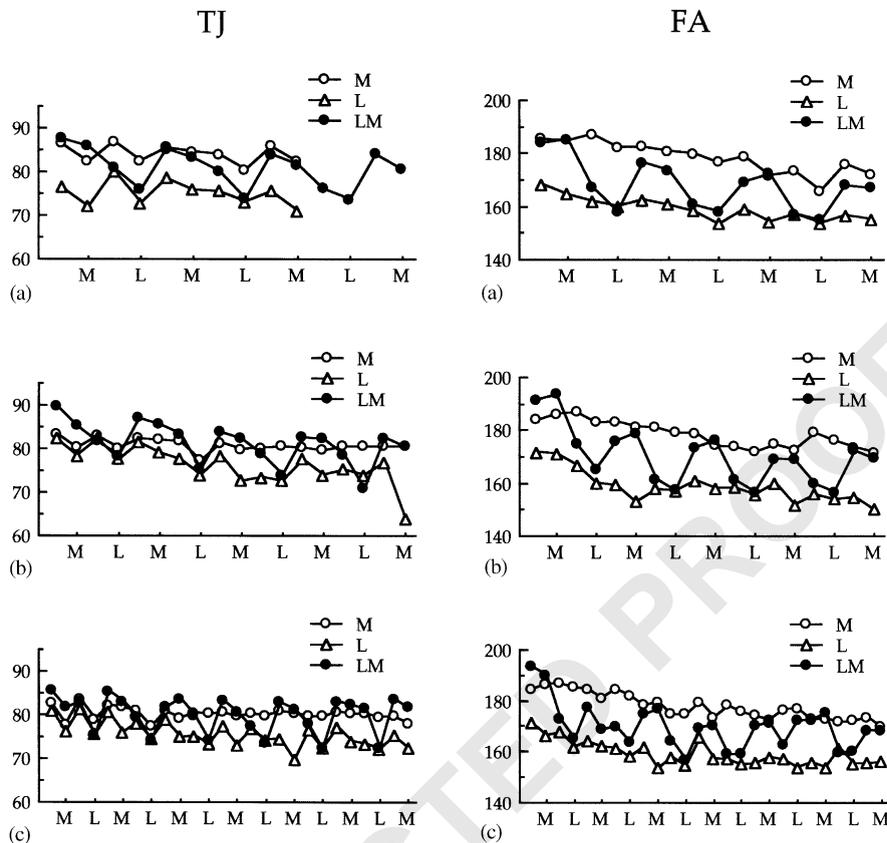


Fig. 8. Superimposed graphs of mixed-LM, all-M and all-L sentences for TJ and FA, showing that L tones do not downstep M tones (values in Hz).

4.3. Discussion

Perhaps the most surprising result of this experiment was the unexpectedly high values of H tones throughout the mixed-HL sentences: the first H tone is set at a value considerably higher than that of its counterpart in all-H sentences, and later “downstepped” H tones do not, in general, drop to values much lower than those of the comparable H tones of all-H sentences. This pattern is unexpected given most earlier accounts of downstep in African languages, in which successively downstepped H tones are described as dropping to values well under the values of comparable nondownstepped H tones (recall the discussion of [Schachter \(1965\)](#) in Section 1).

We attribute this pattern to the interaction of Downstep with H raising, which raises the f_0 value of any H tone to an extra-high value if it is immediately followed by an L tone. H raising affects the first H tone of the sentence most prominently since all subsequent H tones are subject to downstep. Later H tones fail to drop very low since they are also raised before an L. As a result, the descending H tones never cross into the frequency band characterizing M tones (see M tone data in the next section). These interacting effects will be examined more closely in Section 8.

1 5. Experiment 3

3 The next experiment investigates an alternative account of the H raising pattern brought to
 4 light in Experiment 2. This account claims that contrary to appearances, there is no upward shift
 5 of H tones in the mixed-HL sentences at all; rather, the values of H tones in the all-H sentences
 6 with which we have compared them have been shifted globally downward. Such a view is not
 7 implausible given Connell and Ladd's earlier finding that sentence-final H tones in Yoruba
 8 undergo "total downstep," dropping to the level of a preceding L tone (Connell & Ladd, 1990, p.
 9 25). If this principle applied not just to one but to all members of a sequence of sentence-final H
 10 tone syllables (as it does in Kikuyu, see Clements & Ford, 1979), it would produce exactly the
 11 effect just described.

12 A direct way of testing this account is to examine mixed-tone sentences containing H and M
 13 tones. If the global lowering account is correct, the H tones in all-H sentences should be lower
 14 than the comparable nonfinal H tones in mixed-HM sentences, which by hypothesis are not
 15 subject to such lowering.

16 This experiment also tests whether M tones are triggers or targets of downstep in Yoruba.
 17 Courtenay (1971) claimed that M tones are downstepped by L tones in Yoruba. However,
 18 Akinlabi and Liberman (1995) found no downstep (in their terms, "downdrift") of M tones in
 19 MLML sequences, nor of H tones in HMHM sequences.⁷ We will further examine the question
 20 whether H tones are downstepped by L tones "at a distance," that is, across intervening M tones.
 21

22 5.1. Corpus design

23 Three new sets of sentences, mixing M tones with H and L tones in various combinations (LM,
 24 MH, LMH) but otherwise similar in structure to the sentences of Experiment 2, were created to
 25 test these questions. A fourth set consisted of all-M sentences. All these sentences, when pooled
 26 with those of Experiment 2, allow us to compare H, M, and L tones in "like-tone" sentences with
 27 their counterparts in "mixed-tone" sentences containing M tones. (The test sentences are given in
 28 Appendix C.) Recording and analysis were carried out for two speakers (TJ and FA) at the same
 29 time and in the same manner as for Experiment 2.
 30

31 5.2. Results

32 We first examine the like-tone sentences. Fig. 6 shows superimposed f₀ contours of all-H, all-M,
 33 and all-L sentences for each of the three sentence lengths. Graphs are arranged by sentence length
 34 from shortest (a) to longest (c). Each T of the x-axis (to be read as H, M, or L as appropriate) is
 35 aligned with the second measurement point for the syllable bearing it, as in earlier graphs.
 36

37 Somewhat different patterns can be observed for the two speakers. Of TJ's three tone levels
 38 (shown on the left), only L tones show any appreciable decline. This result is in agreement with
 39 earlier studies of Yoruba (see Section 2). In contrast, all of FA's tone levels decline, H tones most
 40

41 ⁷Floating L tones are also sometimes said to downstep following M tones in Yoruba. Laniran (1992, 1995) shows
 42 that for at least some speakers, the major acoustic correlate of such "lowered M tones" is the H raising of the preceding
 43 H tone.

Table 2

Regression analysis of the all-H, all-M, and all-L sentences in Fig. 6 for speakers TJ and FA

Tone	TJ			FA			
	Slope	Intercept	R^2	Slope	Intercept	R^2	
(a)	H	-0.156	93.451	0.217	-1.673	208.506	0.829
	M	-0.259	85.613	0.133	-1.534	188.423	0.802
	L	-0.330	76.987	0.118	-0.753	163.714	0.661
(b)	H	-0.132	90.508	0.190	-1.956	211.013	0.922
	M	-0.108	81.910	0.173	-0.843	185.874	0.735
	L	-0.599	81.983	0.528	-0.766	164.991	0.548
(c)	H	-0.113	91.550	0.268	-1.190	205.195	0.797
	M	-0.057	81.063	0.097	-0.615	186.277	0.822
	L	-0.238	78.865	0.414	-0.452	165.443	0.515

of all. As this result is unexpected, a regression analysis of all data was carried out, as shown in Table 2.

All slopes are significantly different from zero in FA's data, while for TJ only those of the L tones in the mid (b) and long (c) sentences and those of the H tones in the long sentences (c) are significantly different from zero.

Two general observations can be made. First, we see that declination may affect different tone levels at different rates; thus, TJ's L tones decline more rapidly (i.e., show a larger negative slope) than his H and M tones, and FA's H tones decline more rapidly than her M tones, which themselves decay more rapidly than her L tones. Second, the amount of declination along any given tone level is a speaker-specific parameter: thus, FA's H tones decline sharply, while TJ's H tones decline very little if at all. These differences cannot be reduced to the fact that pitch changes are perceived in terms of proportional rather than absolute differences. If we calculate the syllable-by-syllable drop in terms of percentages, we find that over FA's sentences of all lengths, H tones decay by an average of 1.5% per syllable, M tones by 1.1%, and L tones by 0.8%. In contrast, TJ's L tones decay by 1.0% per syllable, and his other tones show no significant decay, apart from the one case noted above.⁸

To test the hypothesis that tones in all-H sentences are globally lowered by total downstep or some similar effect, we may compare the values of H tones in all-H sentences with the values of the counterpart H tones in mixed-MH sentences. Fig. 7 overlays mixed-MH, all-H, and all-M sentences. The "M" and "H" labels along the x -axis identify the M and H tones of the mixed-MH sentences. Graphs are laid out as before.

⁸A further observation concerns the realization of L tones *within* each syllable. The graphs for the all-L sentences in Fig. 6 confirm the claim of Hombert (1977), based on data from two speakers, that the L tone in Yoruba is realized as a low falling tone on syllables not followed by a higher tone.

1 If it were the case that H tone values are globally lowered in all-H sentences, we would expect H
 2 tones to be lower in all-H sentences than in comparable mixed-MH sentences. Fig. 7 shows that
 3 the opposite is true: H tones are generally *higher* in all-H sentences than in the comparable mixed-
 4 MH sentences. These examples (and further ones to be examined below) allow us to reject this
 5 alternative account of the results of Experiment 2. We conclude that an independent factor of H
 6 raising does indeed operate to raise H tones above their expected values in alternating H and L
 7 tone sequences.

8 Let us now consider the interaction between M tones and other tones. Fig. 7 allows us to
 9 address the question whether M tones downstep H tones. Although H tones in mixed-MH
 10 sentences are generally somewhat lower than H tones of all-H sentences, at least in the shorter
 11 sentences (a, b), this lowering does not increase cumulatively across the sentence in TJ's data, and
 12 so cannot be regarded as downstep as characterized in Section 1. We interpret it instead as a local,
 13 incomplete assimilation of H tones to M tones, perhaps as a result of articulatory undershoot.
 14 Although H tones do lower progressively in FA's mixed-MH sentences, this lowering parallels
 15 that of the H tones in her all-H sentences. It is therefore not due to downstep operating on H
 16 tones. We suggest that it can be understood as an effect of downdrift (or tone-specific declination)
 17 operating "in the background" across the H tone reference line, gradually lowering the value at
 18 which H tones can be realized even when other tones intervene. ("Background downdrift" is
 19 further discussed just below.)

20 Let us next consider whether L tones downstep M tones, as claimed by Courtenay (1971). Fig. 8
 21 overlays mixed-LM sentences with all-M and all-L sentences. The "M" and "L" labels along the
 22 x-axis identify the M and L tones in the mixed-LM sentences.

23 If L tones downstepped M tones, we would expect M tones to have progressively lower values
 24 in LM sentences than the comparable M tones in the all-M sentences. Fig. 8 shows that this is not
 25 the case for either speaker: we see that the rate of M tone decline is about the same in both
 26 sentence types. This pattern is different from that shown by the downstepping H tones in the
 27 mixed-HL sentences of Experiment 2, which descended more rapidly than the comparable H tones
 28 in all-H sentence (Fig. 5).⁹ For these reasons, we conclude that L tones do not downstep M tones
 29 in these data.

30 We further find that minimal values of L tones in mixed-LM sentences show the same pattern of
 31 gradual decline that they do in all-L sentences. Since the decline is similar in both sentence types,
 32 we cannot attribute it to the presence of the M tones. We interpret it instead as an effect of
 33 "background" downdrift operating across the L-tone reference line, gradually lowering the value
 34 at which an L tone can be realized even when M tones intervene. We saw a similar effect of
 35 background downdrift operating on H tones just above, and will see further evidence for
 36 background downdrift on L tones in the later discussion.

37 Let us finally consider whether L tones can downstep H tones "at a distance" across intervening
 38 M tones. Fig. 9 superimposes mixed-LMH sentences upon all-H sentences, all-M sentences, and
 39 all-L sentences for two length conditions, medium (a) and long (b). (Shorter sentences are
 40 unrevealing in this case.) The x-axis is labelled with the H, M and L tones of the mixed-LMH
 41 sentences.

42 ⁹The M tones in TJ's longer mixed-LM sentences are generally *higher* than the comparable M tones in all-M
 43 sentences, suggesting that this speaker may employ a dissimilatory principle of M raising in the context of L tones.

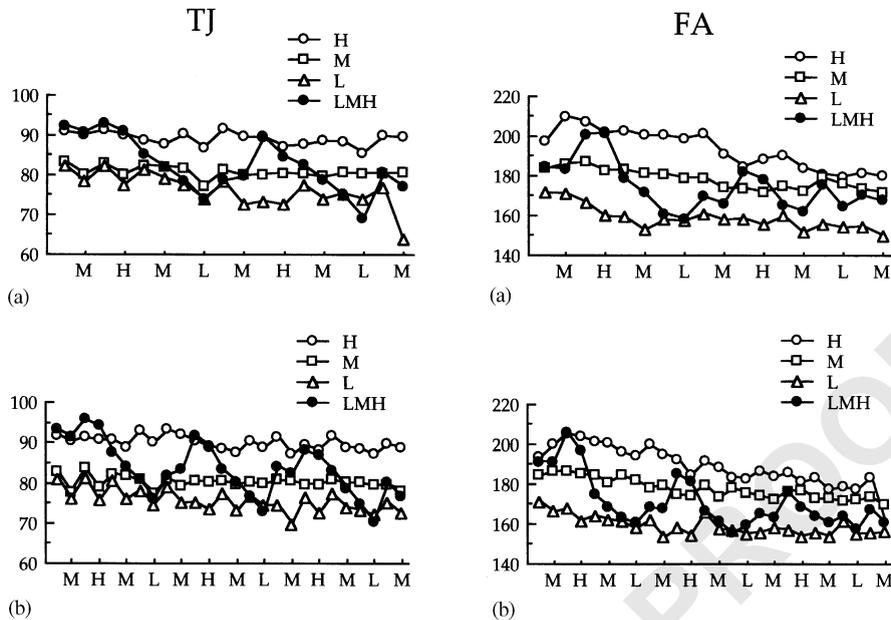


Fig. 9. Superimposed graphs of mixed-LMH, all-H, all-M, and all-L sentences in medium-length (a) and long (b) sentences, for speakers TJ and FA. Only one graph (lower right) shows apparent downstep of H tones after LM sequences (values in Hz).

All displayed portions of the mixed-LMH sentences start and end with M tones (as before, we omit the final carrier frames). In general, values of H, M and L tones in the mixed-LMH sentences tend to approach those of the comparable tones in the like-tone sentences fairly closely, and we do not find the consistent upward shift of early H tones to extra-high levels that we found in the mixed-HL sentences (cf. Fig. 5), nor any consistent downstepping effect on later H tones. However, we note some interspeaker differences. First, TJ's first M tone is much higher than expected in both sentence lengths, perhaps as an effect of local assimilation to the following H. Second, in both of FA's sentence lengths, H and M tones are lower than the corresponding tones of the all-H and all-M sentences, and in her longer LMH sentence (b) H tones descend more rapidly than they do in the corresponding all-H sentence. These effects might suggest that L tones downstep both H and M tones here, but we have already seen (in Fig. 8) that L tones do not downstep M tones in other data produced by this speaker. As speaker TJ shows no similar effects, we must conclude that any such effects, if they can be confirmed, are speaker-specific and perhaps context-specific at best.

These graphs also show that L tones tend to have the same pattern of gradual decline in the mixed-LMH sentences that they have in all-L sentences, giving further evidence that L tones are subject to background downdrift, even when separated by other tones.¹⁰

¹⁰This effect is clearly visible in TJ's data, less so in FA's. We have no explanation for the unexpectedly high value of FA's penultimate L syllable; especially in (a), this syllable appears to bear a phonological falling (HL) tone.

1 5.3. Discussion

3 The presence of background downdrift across L tones as observed in Figs. 8 and 9 sheds new
 4 light on the question of whether L tones undergo downstep in the mixed-HL sentences shown
 5 earlier in Figs. 2–4. While we found that L-tone values decline in these sentences from the third L
 6 tone onward, Fig. 5 showed that declining L tones do not reach values lower than those of
 7 comparable L tones in all-L sentences. We conclude that there is no clear evidence that L tones in
 8 mixed-HL sentences undergo downstep; instead, they appear to undergo background downdrift,
 9 just as they do in mixed-tone sentences in which downstep does not apply (Figs. 8 and 9).

10 The main conclusions of Experiments 1–3 can be summarized as follows. Downstep applies to
 11 H tones, but not to L tones, in mixed-HL sentences. H raising interacts with downstep to raise the
 12 values of all H tones preceding L tones. L tones do not downstep M tones, nor do M tones
 13 downstep H tones. Moreover, L tones do not appear to downstep H tones at a distance, but only
 14 when adjacent to them. Finally, there is no evidence for global lowering (“total downstep”) of all-
 15 H sentences.

17

19

21 6. Experiment 4

22 The experiments described so far have revealed relatively little evidence of “preplanning” or
 23 look-ahead effects in Yoruba tone production. However, a significant effect of this sort was found
 24 in TJ’s L-initial sentences of Experiment 1. Accordingly, in order to study possible preplanning
 25 effects more closely, we conducted a further experiment, described in this section.¹¹ The
 26 experimental data will also allow us to further test (and ultimately confirm) the independence of
 27 downstep and H raising.

29

31

33 6.1. Corpus design

34 Three sets of test sentences were devised to determine whether the presence of an L tone later in
 35 a sentence can cause the anticipatory raising of an earlier H tone “at a distance,” that is, several
 36 syllables away. The initial four tones in these sentences satisfy all possible expansions of the
 37 formula “HXHY” obtained by substituting H, M, and L tones for the variables *X* and *Y*, as
 38 shown in (4). To eliminate any final lowering effects, these tone sequences were embedded in a
 39 frame bearing the following tone sequence HLHHMH. All sentences were recorded by speakers
 40 TJ, KG, FA, and BJ (the test sentences are listed in Appendix D):

41

43

¹¹ An experiment based on a smaller number of speakers was described in Laniran (1992) and Laniran and Clements (1995). We point out here that the published version of the latter paper excised the following text at the bottom of p. 736, column 1: “examples of graphs (c) and (f) with the” and the following text at the bottom of column 2 of the same page: “triangles, showing the H_1LH_2L se-”.

	Set A (HHHY...)	Set B (HMHY...)	Set C (HLHY...)	
1				
3	1. HHHH...	4. HMHH...	7. HLHH...	(4)
	2. HHHM...	5. HMHM...	8. HLHM...	
5	3. HHHL...	6. HMHL...	9. HLHL...	

7 Sentences 7–9 provide contexts for downstep, and sentences 3 and 6–9 provide contexts for H raising (two contexts, in the case of sentence 9).

9 The null hypothesis is that there are no preplanning effects in Yoruba tone production. If this hypothesis is true, values of the first H tone in these sentences should be independent of the choice of the fourth tone. Thus, the first H tone in sentence 9 should have the same peak value as the first H tone in sentences 7 and 8, since the presence of the later L tone in sentence 9 should have no effect on its realization. A further prediction (abstracting from any further effects such as declination) is that the initial H tone in sentence 9 should have the same peak value as the H that precedes the L in sentences 3 and 6, since all occur in the H raising context, and none are subject to downstep.¹²

17 6.2. Results

19 Data for all speakers are shown graphically in Fig. 10.

21 Graphs in the first column present the overlaid f₀ tracks of the three sentences of Set A, those in the second column those of Set B, and those in the third column those of Set C. The first H of the following sentence frame is included at the end of each graph. Within each graph, the only factor that varies is the choice of the fourth tone, *Y*, of the formula HXHY. Two f₀ values are shown for each syllable. For ease of reference, the first three f₀ values have been labelled a, b, and c. Points a and b always correspond to the first H tone. Point c also corresponds to the first H tone in the set C sentences, due to the effect of Tone Spread (2) onto the following L tone syllable. In all other cases, point c represents the first value of the lexical tone of the second syllable.

29 These graphs show that all speakers have applied Downstep to the second tone in HLH sequences as expected (cf. the first and second H tone peaks in the set C sentences). An examination of sentences 3, 6, and 9 (tracked with black circles) shows that all speakers have also applied H raising: H tones are in all cases higher before L tones than they are in the comparable sentences (tracked with white circles) where they occur before H and M tones. This result confirms the independent role of H raising in Yoruba.

35 Let us now attempt to determine whether there is a “long-distance” dependency between the peak values of the first H tone and the choice of tone *Y* in the set C sentences. As noted above, the null hypothesis is that there will be no such effect. We thus expect to find no difference between the peak value of the first H tone in sentence 9 (HLHL) and those of sentences 7 (HLHH) and 8 (HLHM). We see, however, that for three speakers, TJ, KG, and BJ, the values of the first H at points b and c in sentence 9 are higher than those at the same point of sentences 7 and 8. In other words, for these speakers, the choice of *Y*=L in the HLHY sentences correlates with extra-high values of the first H tone, a long-distance effect. Quantitative values for these speakers, showing

¹²We thank Mary Beckman for pointing out this further prediction to us.

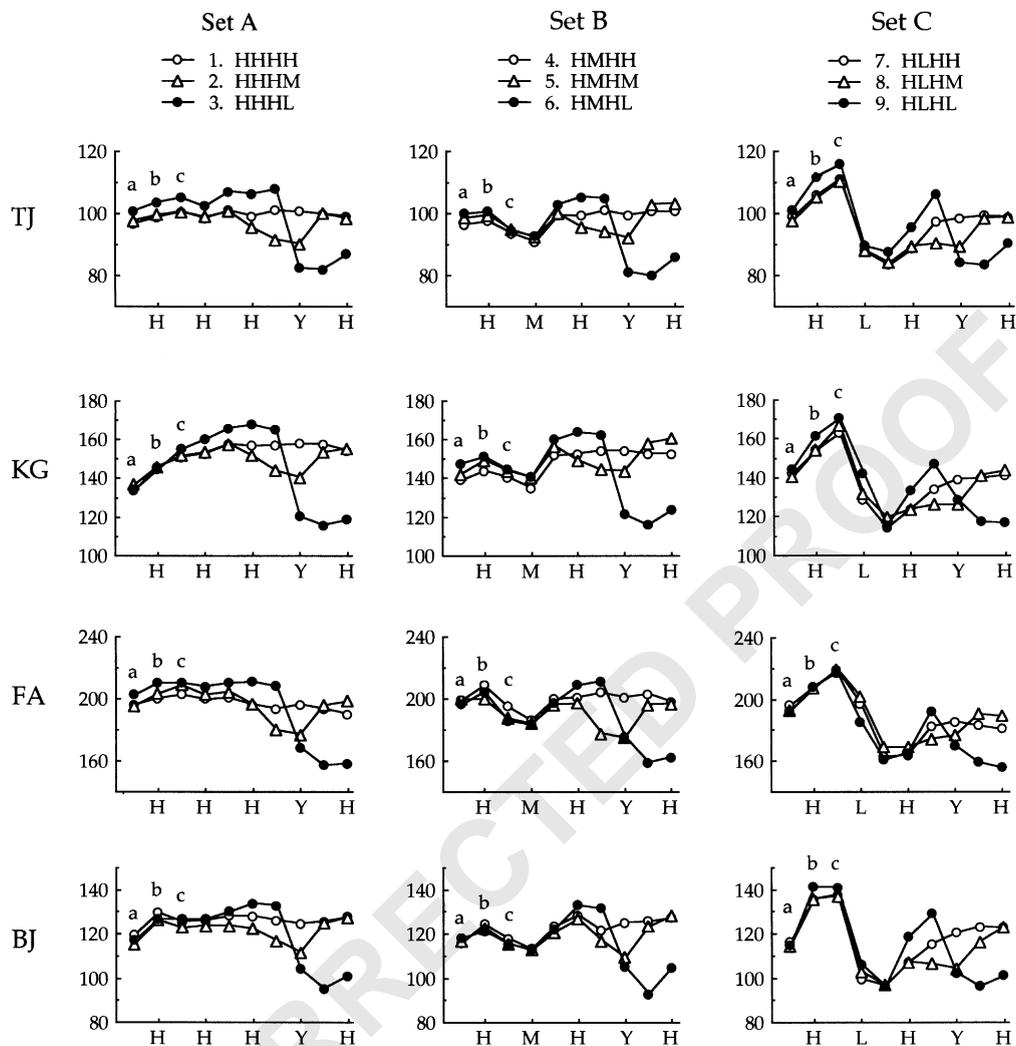


Fig. 10. Realizations of HXHY sentences for four speakers (Experiment 4) (values in Hz).

differences between the first H peak in sentence 9 and the comparable peaks in sentences 7 and 8, are given in the second and third columns of Table 3.

A second prediction of the null hypothesis is that the first H tone in sentence 9 (HLHL) should have the same peak values as the last H tones in sentences 3 (HHHL) and 6 (HMHL), since none of these tones are downstepped. This prediction is also incorrect. The peak value of the first H tone in the HLHL sentence is consistently higher, this time for all four speakers and by somewhat larger margins of 5–9 Hz, than the peak H tone values of the last H tones in sentences 3 and 6. Values are given in the fourth and fifth columns of Table 3. This result again demonstrates a “long-distance” anticipation of the L tone three syllables away.

As the f_0 differences here are rather small, averaging around 6 Hz, we performed a two-way ANOVA on these data to detect interactions of the first H tone (H_1) with tone Y as well as with the immediately following tone X in HXHY sentences at points a, b, and c. An interaction of H_1

1 Table 3
Long-distance effects on H tone realizations (Experiment 4)

3 Speaker	Δ (9–7)	Δ (9–8)	Δ (9–3)	Δ (9–6)	Mean
5 TJ	4.6	5.3	7.8	10.5	7.1
5 KG	7.5	3.5	2.8	6.5	5.1
5 BJ	3.7	4.1	7.3	8.1	5.8

7 The four middle columns show differences in Hz between the average peak value of the first H tone in sentence 9 and
9 those of the H tone peaks in sentences 7, 8, 3, and 6, for three speakers.

11 Table 4
ANOVA results, showing p -values of the interaction of the first H tone with tones X and Y in HXHY sentences at
13 points a, b, and c for two speakers (Experiment 4)

13 H tone at	TJ	KG
15 Point a: X	n.s.	n.s.
15 Y	0.0140	0.0191
17 Point b: X	0.0001	n.s.
17 Y	0.0005	n.s.
19 Point c: X	0.0001	0.0001
21 Y	0.0065	n.s.

23 with tone X would confirm a local effect of H raising, and an interaction of H_1 with the distant
25 tone Y would reveal a long-distance anticipatory effect. The null hypothesis is that the f_0 values of
 H_1 are independent of both X and Y . p values under 0.05 are taken as showing that the null
hypothesis should be rejected.

27 Significant effects were found for two speakers, TJ and KG, as shown in Table 4.

29 As expected, there was a significant correlation between H_1 and the immediately following tone,
31 X , for both speakers. This confirms the local application of H raising. As for the long-distance
case of special interest here, Table 4 shows that the identity of tone Y significantly affected the
33 value of H_1 at point a for both speakers, and also at points b and c for TJ. In other words, for
these two speakers, the raising of the initial H in sentence 9 is significantly correlated with the
presence of an L tone three syllables away.¹³

35 This result can be explained on the view that some speakers make room for the second
downstep in HLHLH sequences by raising the pitch of the first H to an extra-high value, thereby
37 expanding the available pitch space. This result confirms the existence of small but significant
preplanning effects in Yoruba tone production. These results do not generalize to all of our data,
39 however, since Experiment 1 found no significant anticipatory effects in similar sentences for two
other speakers, nor for TJ in the H-initial sentences.

41 As a final point, the data in this experiment allow us to confirm that downstep does not apply to
L tones. Table 5 compares average L-tone minima in the three contexts provided by sentences 3
(HHHL), 6 (HMHL), and 9 (HLHL). (Recall that all these sequences are followed by a H tone in
43

¹³There was no significant interaction between tones X and Y for any speaker.

1 Table 5
 2 Average L-tone minima (in Hz) of the final L tone in sentences 3 (HHHL), 6 (HMHL), and 9 (HLHL), for four
 3 speakers

Speaker	3: HHHL	6: HMHL	9: HLHL
5 TJ	83.0	80.1	83.5
6 KG	115.8	116.2	117.1
7 FA	157.3	159.3	156.2
8 BJ	95.3	93.0	96.5

9 the sentence frame.) Only sentence 9 contains the HL downstep trigger. If L tones undergo
 11 downstep after H tones, the value of the second L tone in this sentence should be lower than that
 of the comparable L tones in sentences 3 and 6.

13 We see that this prediction is not confirmed. For one speaker (FA), the second L of sentence 9 is
 15 lower than the comparable L's of sentences 3 and 6 by only 1–3 Hz, a difference much smaller
 17 than what could be attributed to downstep. For the others, the second L is actually higher than
 the others (though again by a very small amount). Thus, downstep does not apply to L tones in
 these data.

19

21 7. Alternative interpretations

23 The data observed up to now have suggested that both downstep and H raising operate on H
 25 tones in sequences of alternating H and L tones. These two principles appear to interact to raise
 27 downstepping H tones to values higher than what would be expected from the operation of
 29 downstep alone. Before concluding that this interpretation is correct, however, we must consider
 31 two alternative accounts of our data, both of which deny that there is any interaction between
 downstep and H raising. The first denies that downstep applies in Yoruba at all, attributing the
 apparent downstep pattern to the local assimilation of H tones to preceding L tones. The second
 admits that both downstep and H raising are at work, but claims that they apply in strictly
 complementary contexts in which they never interact. We shall show that neither of these accounts
 is consistent with the data examined up to now.

33 7.1. Downstep reinterpreted as local assimilation

35 Under the first account, the downstep effects observed in Experiment 1 are attributed to local
 37 assimilation of the second H tone to the flanking L tones in the context HLHL. It will be recalled
 39 that in all Experiment 1 sentences, only *single* H tones intervened between L tones. Lowering of
 41 single H tones flanked by L tones could result from local assimilation or articulatory undershoot.
 Since, moreover, L tones undergo background downdrift in our data, local assimilation of H
 tones to L tones produced along a falling ramp could be expected to produce a declining H tone
 pattern similar to what we have observed.¹⁴

43 ¹⁴We thank Bruce Connell for bringing this interpretation to our attention, as well as for useful suggestions as on
 how it can be tested.

1 Several observations show, however, that such an account cannot be maintained for Yoruba. If
 2 it were correct, we would expect the first H in L-initial sentences to be lower than the first H in H-
 3 initial sentences, due to assimilation to the preceding L tone. As we have seen, this is not the case;
 4 of the speakers examined in experiment 1, only FA showed a trend in this direction, and a very
 5 small one at that. A local assimilation account would also predict that H tones in mixed-HL
 6 sentences should decline at about the same rate as L tones to which they are keyed; but inspecting
 7 the initial LHLH sequence in Figs. 2–5, we see that H tones drop at a much faster rate than L
 8 tones. Furthermore, as noted earlier, for two speakers (TJ, KG) the first L tone in L-initial
 9 sentences is considerably lower than the second. If H tones are keyed to the preceding L, the local-
 10 assimilation account would predict that the first H tone will be lower than the second, when in fact
 11 it is much higher.

12 A very direct way of discriminating between the two accounts is to examine sentences in which
 13 not just single H tones, but three or more H tones follow L tones, as in the sequence HLHHH.
 14 Laniran (1992) has previously reported local assimilation effects on the first one or two H tones in
 15 such sequences. If only local assimilation is at work in mixed-HL sentences, as the alternative
 16 account claims, the first and perhaps second H tone in a H-tone sequence following the L should
 17 be lower than the initial H tone, but the third and any subsequent H tones in the sequence should
 18 be as high as the initial H, since assimilation, as a local effect, should not affect them. If instead
 19 downstep is at work, as has been maintained up to now, all H tones following the L, including
 20 crucially the third and any later ones, should be lower than the initial H.

21 Relevant data are provided by sentence 7 of Experiment 4, which was displayed earlier in
 22 Fig. 10. Productions of this sentence are redisplayed for all four speakers in Fig. 11.

23 The sentence shown here, *Máyòmi rélá*, bears the tone sequence HLHHH. (This sequence was
 24 followed by the sequence LHHMH in the carrier sentence.) All speakers show some degree of
 25 local assimilation on the first H tone following the L, as shown by point a, which is lower than all
 26 following H tone values. While this assimilation perseveres into the second H tone syllable in KG
 27 and BJ's traces (represented by points b and c), for all speakers it essentially ceases to operate by
 28 the beginning of the third H tone syllable (point d), whose pitch nevertheless remains well below
 29 the peak value of the sentence-initial H tone. (This is true in spite of the fact that the third H tone
 30 of the H sequence is subject to H raising by the following L tone of the carrier sentence.) The
 31 entire H tone sequence must therefore be regarded as downstepped. We may conclude that for all
 32 speakers, both local assimilation and downstep are in operation here. These processes operate
 33 together on the first one or two H syllables following the L, but only downstep operates
 34 thereafter.¹⁵

37 7.2. Downstep and H raising reinterpreted as noninteracting principles

38 According to the second account, H raising applies just once, to the sentence-initial H tone.
 39 Since downstep applies only to noninitial H tones, the two principles do not interact. A model for
 40 this has been motivated by recent work by Rialland and Somé (2000). These writers show that
 41

42 ¹⁵It remains to be determined how much (if any) of the lowering effect in the mixed-HL sentences of experiment 1 can
 43 be attributed to local assimilation. This question could be answered by examining sentences of the form HLH*LH*...,
 in which each H* consists of three or more H tones. No examples of this structure were examined in Laniran (1992).

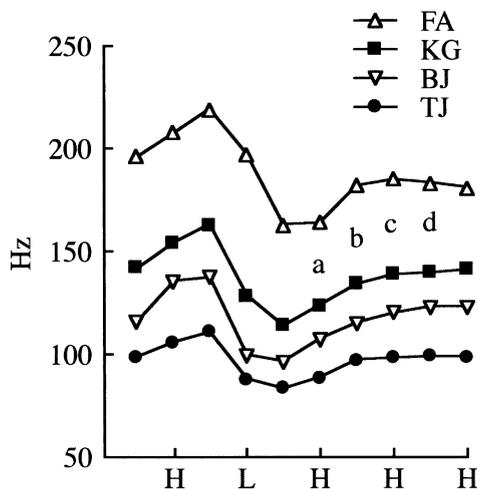


Fig. 11. Realizations of HLHHH sentences for four speakers, showing both carry-over assimilation and downstep effects on the final three H tones.

Dagara speakers make room for later downsteps in a sentence by raising the pitch of the first H tone to a value roughly proportionate to the number of following downsteps. Downsteps of the following H's in the sequence are then keyed to the value of this first H. This principle, which might be called "regressive upstep" as it has the effect of upstepping H tones in a cumulative manner as we proceed from right to left across the sentence,¹⁶ yields an effect superficially similar to that of H raising as we have characterized it above.

This principle predicts very robust long-distance anticipation effects. As such effects are relatively small and speaker-specific in Yoruba, as we saw in the discussion of Experiment 4, we shall consider a restricted version of this model in which H raising shifts the H tone of the first HL sequence to a fixed extra-high value, regardless of how many downsteps follow. This extra-high value then establishes the ceiling with respect to which the next downstepped H tone is keyed. This value in turn sets the ceiling for the next downstep, and so forth across the downstep span. In this model, once the first H tone has been initialized, values of each succeeding H tone are determined only by the values of preceding H tones. H raising does not apply to noninitial H tones. As a result, it never applies to a downstepped H tone, and so the two principles do not interact.

To see whether this model can be applied to Yoruba data, we may compare sentence 8 of Experiment 4, containing the initial sequence HLHM, with sentence 9 of the same experiment, containing the initial sequence HLHL. Under both accounts, H raising will apply to the first H tone and downstep to the second in both sentences. What distinguishes them is that under the first account, in which H raising applies to *all* H tones preceding L tones, H raising also applies to the second H of sentence 9, while under the alternative it cannot. Thus, if the original account is correct, the second H of sentence 9 should be higher than that of sentence 8. Under the alternative account, both should have about the same value, abstracting from any small differences that might result from the anticipatory raising strategy observed for the two speakers TJ and KG.

¹⁶A regressive upstep model was tentatively suggested for Yoruba in an earlier publication (Laniran & Clements, 1995). For reasons to be discussed just below, we no longer hold that Yoruba tonal data can be viewed in this way.

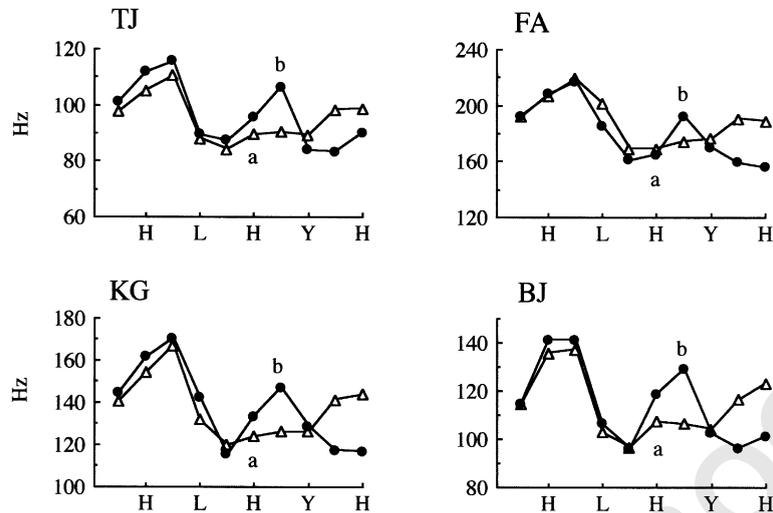


Fig. 12. Realizations of sentences 8 (HLHM) and 9 (HLHL) for four speakers, allowing comparison of the peak value of the second H tone (points a and b) in each. HLHM sentences are tracked with white triangles, HLHL sentences with black circles. It can be observed that H raising affects the second H tone of the HLHL sentence at point b.

Fig. 12 compares sentence 8 (HLHM) and sentence 9 (HLHL) for all four speakers.

In all cases, the second H tone in the HLHL sentence (tracked by black circles, whose peak is located at point b) is higher than the one in the HLHM sentence (tracked by white triangles, whose peak is located at point a) by a large margin, greatly exceeding the small differences that might be attributed to the pervasive effects of long-distance anticipatory raising. This is true most crucially for FA, who employs no anticipatory raising at all in these data. This striking difference must be attributed to the separate operation of H raising on the second H tone. It confirms that H raising interacts with downstep on noninitial tones for all speakers.

8. The interaction of downstep and H raising

Our discussion has confirmed the existence of two interacting principles in Yoruba tone production, downstep and H raising, but has not so far raised the obvious question: why should both processes coexist in a tone language like Yoruba? Or alternatively, why should one exist, given the other? This section will address these questions by examining quantitative aspects of tone production in Yoruba. It will show that although Yoruba speakers implement downstep and H raising by quantitatively different means, their realization strategies “conspire” to insure that downstepping H tone will not penetrate the frequency band reserved for M tones.

The following subsections consider quantitative aspects of downstep and H raising, as well as of two further principles that interact with them, resetting and L assimilation.

8.1. Downstep

Earlier work on the modelling of downstep falls into two traditions, which we will term “Categorical” and “Gradient” for convenience. While Categorical models differ in detail (see e.g.

1 Stewart (1993) for a recent example), most have agreed on a number of basic assumptions.
 2 Sentence-initial H tones are typically set to an abstract initial value of 1. If no downstep occurs
 3 anywhere in the following tone sequence, all H tones will be assigned this integer value, and will
 4 theoretically be realized at the same pitch. If a subsequent H tone is downstepped, it is assigned an
 5 additional value of 1, representing a pitch drop of one unit. Each further downstepped H tone in
 6 the downstep span adds 1 again, so that the pitch of any H tone is determined by the sum of the
 7 integers that have been assigned to it. For example, the three H tones in the downstepping
 8 sequence HLHLH receive the integers 1, 2, 3, respectively. By assigning tones integers instead of
 9 real numbers, this approach claims that there is a discrete though theoretically open-ended
 10 number of levels to which H tones can be assigned. It also predicts that the contrast between
 11 phonologically different tones (such as H and M) may be completely neutralized if they receive the
 12 same integer value in the same context (see Snider, 1998).

13 The Categorical approach as just summarized says nothing about how integers are interpreted in
 14 terms of actual f0 contours. This question is addressed in Gradient models of downstep. One
 15 influential model of this type, proposed by Liberman and Pierrehumbert (1984), defines
 16 downstepping patterns as a gradual decay toward an abstract reference line, or asymptote. In this
 17 approach, the value of any H tone in a downstepping sequence, H_n is given by

$$18 \quad H_n = d(H_{n-1} - r) + r \quad (5)$$

19 in which H_{n-1} is the value of the preceding H tone, d is the downstep coefficient (a ratio between 0
 20 and 1), and r is the value of the reference line towards which the H tones decline. This expression
 21 assigns f0 values to downstepping H tones from left to right within a downstep span. Thus, for
 22 example, if the first H tone in the sequence HLHLH is initialized at 200 Hz, d is set at 0.6, and r
 23 is set at 100 Hz, the second H will have a value of 160 Hz, the third will have a value of 136 Hz, and
 24 so forth. This method of f0 assignment describes an exponentially decaying curve in which each
 25 step down is proportionally identical to the preceding one in terms of its distance from the
 26 reference line r . Later downstep intervals are progressively smaller than earlier ones, and tend to
 27 become vanishingly small as the reference line is approached. This approach could be called a
 28 “soft-landing” model of downstep implementation as it describes a curve similar to that of an
 29 aircraft gliding smoothly down to a landing strip.

30 A limitation of this approach is that it does not provide a speaker-independent characterization
 31 of downstep registers for given utterances, since values of d and r typically vary from speaker to
 32 speaker. However, as several writers have observed (Clements, 1990; Ladd, 1990, appendix;
 33 Liberman, Schultz, Hong, & Okeke, 1993, p. 156), the two approaches are complementary in the
 34 sense that the Gradient model can be regarded as providing speaker-specific f0 interpretations of
 35 the integer sequences provided by the Categorical model. The fundamental insight is that the
 36 integer value assigned to any H tone by the Categorical model represents the number of times
 37 downstep has applied to it, plus 1. Thus downstep has not applied at all to the initial H tone in the
 38 downstepping sequence, whose integer value is 1, downstep has applied once to the first
 39 downstepped H tone, whose integer value is 2, and so forth. If we thus subtract 1 from the integer
 40 n assigned to any H tone, we obtain the number of times it has been downstepped. To obtain the
 41 height of any downstepped tone H_n above the reference line, therefore, we can assign $n - 1$ as an
 42 exponent of the downstep factor d and multiply the result, d^{n-1} , by the distance of the initial H
 43 tone (H_1) above the reference line, $H_1 - r$. To obtain the actual value of the H tone H_n , we restore

Table 6

Values of downstepping H-tone peaks 1–4 in the first downstep span of the L- and H-initial sentences of Experiment 1 for speakers TJ, FA, and KG

	TJ			FA			KG		
	L-in	H-in	Pooled	L-in	H-in	Pooled	L-in	H-in	Pooled
H_1	102.1	100.7	101.4	211.7	219.9	215.8	149.4	146.7	148.0
H_2	94.6	92.1	93.4	182.6	189.0	185.8	127.1	123.7	125.4
H_3	88.3	85.2	87.3	174.1	175.8	175.0	104.4	101.7	103.1
H_4	84.6	82.8	83.7	170.4	174.5	172.4	—	—	—
H_f	77.4	77.7	77.6	168.1	170.5	169.3	84.2	81.7	82.9

Values are averages in Hz over mean values of sentences of all lengths. The last row shows average values of sentence-final H tones (H_f).

r to the result, as shown:

$$H_n = d^{n-1}(H_1 - r) + r. \quad (6)$$

As an example, given the numerical values of the example following (5) above, the third H tone in a downstepping HLHLH sequence would be assigned the integer 3 under the Categorical model and would then receive the speaker-specific value of 136 Hz ($0.6^2(200 - 100) + 100$) under the Gradient model.¹⁷

Let us now consider the quantitative scaling of downstepping tones in Yoruba, asking the following questions: Do speakers use the soft-landing algorithm described in (6)? If not, what other algorithms are used? We shall see evidence below that while one speaker does use the soft-landing approach to downstep implementation, others use a “hard-landing” model in which downstepped H tones fall abruptly to the reference line in one or two steps.

Let us consider the downstep patterns of speakers TJ, KG, and FA. Values of the downstepping H tone peaks $H_1 - H_4$ in the first downstep span for these speakers are given in Table 6 for the L-initial and H-initial sentences of Experiment 1, together with values of the final H tone (H_f). All values are averages in Hz over mean values of sentences of all lengths. Since these values are taken only within the first downstep span, they exclude values of reset tones; for example, a value of H_4 was excluded from the average if it was higher than that of H_3 .

An examination of Table 6 reveals two general patterns. For TJ and FA, the magnitude of the drop between two successive H tones decreases steadily across the downstep span. For KG, in contrast, the drops are about equal in magnitude. (Only three values are given for this speaker, who normally reset his H tones on H_4 .)

Let us attempt to see what principles underlie these patterns. For TJ, the values of the first four downstepped H tones in the first span can be modelled to a good first approximation by the soft-landing model embodied in expression (6). We can take his value of the reference line r to be equal to that of H_f , the sentence-final H tone. We thus have $r = 77.4$ Hz for the L-initial sentences and $r = 77.7$ Hz for the H-initial sentences. The value of d is then calculated on the basis of H_1 , H_2 and r by

¹⁷ Expression (5) assigns a Hz value to any H tone regardless of the sequence in which values are calculated, and so unlike the algorithm in (4) does not require left-to-right iterative application.

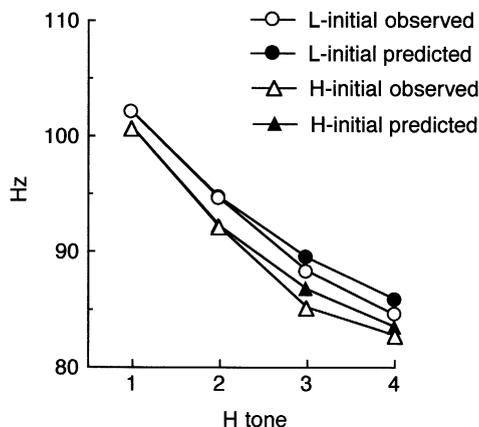


Fig. 13. Plot of observed vs. predicted H tone values in the initial downstep span of the L- and H-initial mixed-HL sentences for speaker TJ, as predicted by the expression $H_n = d(H_{n-1} - r) + r$ (values of r and d for each data set are given in the text).

$$d = H_2 - r/H_1 - r. \quad (7)$$

Given the values of H_1 and H_2 shown in Table 6, we obtain $d = 0.70$ for the L-initial sentences and $d = 0.63$ for the H-initial sentences.

We may test the fit of this model to the data by predicting the values of later H tones from those of H_1 by expression (6) and comparing them with the observed values. The result is shown separately for L- and H-initial sentences in Fig. 13.

Fig. 13 overlays the observed values H tone values (shown by the white tracking symbols) and the predicted H tone values (shown by the black symbols). As this figure shows, the fit between observed and predicted values of the later tones¹⁸ is very good; the largest difference between predicted and observed values is 1.3 Hz, which is about what we would expect if H downdrift is operating in the background (cf. Table 2). This fit confirms the choice of H_f as establishing the reference line r . This choice has an obvious functional motivation, as it allows downstep to apply freely across sequences of H tones while keeping them well above the values of M tones.

The “soft-landing” model was also fitted to the independent data provided by TJ’s mixed-HL sentences of Experiment 2 (displayed in Fig. 5). As we did not record values of the final H tones of the carrier sentence in these sentences, we took $r = 77.4$ Hz, as in the L-initial sentences of Experiment 1. Proceeding on this basis, we obtained a downstep ratio of 0.68. Comparison of predicted and observed values for H_3 and H_4 gave the following result: H_3 predicted = 88.7, observed = 87.5 (−1.2 Hz); H_4 predicted = 85.1, observed = 85.8 (+0.7 Hz). These small discrepancies are reasonable considering that Experiment 2 data set (containing three mixed HL-sentences) was smaller than that of Experiment 1 (with eight sentences). This result provides independent confirmation of the appropriateness of the soft-landing model to TJ’s data.

FA’s Table 6 pattern appears at first sight to lend itself to a similar analysis, but we must factor out her rather large amount of H tone downdrift, as shown earlier in Fig. 6 and Table 2, which

¹⁸Comparison of the H_1 and H_2 values, which are identical ($H_1 = 101.5$, $H_2 = 93.9$) is irrelevant, since these values were used in calculating the downstep ratio.

1 accounts for a large part of the negative f_0 slope here. Estimating the background downdrift rate
 2 at about 4 Hz per syllable in these mixed-HL sentences, FA's adjusted values for H_2 and H_3 are as
 3 follows: for H_2 , L-initial = 190.6 Hz, H-initial = 197.0 Hz, and for H_3 , L-initial = 190.1 Hz, H-
 4 initial = 191.8 Hz. Given these adjusted values, downstep between H_2 and H_3 is negligible in the L-
 5 initial sentences, and much smaller than first appeared in the H-initial sentences. Moreover, the
 6 adjusted H_4 values are now higher than the H_3 values (L-initial = 194.4 Hz, H-initial = 198.5 Hz),
 7 and so H_4 is not downstepped but reset with respect to H_3 . Application of the "soft-landing"
 8 model to these adjusted figures gives a poor fit to the observed data. We conclude that this model
 9 does not fit FA's downstep pattern. It can be better described by a "hard-landing" model in which
 10 the first downstepped H tone drops once by about 20–23 Hz and subsequent H tones decline by
 11 the H tone downdrift factor of about 3–4 Hz per syllable.

Nor can speaker KG's data be modelled in terms of the soft-landing model. In his data, H_2 and
 12 H_3 drop in equal steps of about 22.5 Hz to a value of about 103 Hz, well above his values for H_f .
 13 In most of his sentences, resetting applies immediately thereafter. This strategy is defined by

$$14 \quad H_n = H_{n-1} - d \quad (n = 2 \text{ or } 3), \quad (8)$$

15 where d is now an absolute value in Hz rather than a ratio. Although different from the previous
 16 patterns, this approach also avoids any possible confusion of downstepping H tones with M tones.

17 To what extent do the speaker-specific values observed in the initial downstep spans, as shown
 18 in Table 6, generalize to later spans? For TJ and KG,¹⁹ the downstep values observed in the initial
 19 span do not generalize to the second span. TJ's later downstepped H tones nearly always drop by
 20 at least 5 Hz, a much greater amount than would be predicted by the downstep ratio of around 2/3
 21 employed in the initial span. This difference could be explained on the view that in order for
 22 downstep to be perceived in later parts of longer sentences, it must have a value of at least 5 Hz or
 23 so; a downstep ratio of 2/3, if maintained later in the sentence, would yield an average drop of
 24 4.5 Hz, which, given the expected variability around the mean, would yield many token-specific
 25 drops well under this threshold value. In contrast to TJ, KG's later downstepped H tones drop by
 26 an average of 10.6 Hz, a value well *under* the 22.5 Hz registered in the initial span. For this
 27 speaker, continued use of the large absolute downstep value of 22.5 Hz in later spans would take
 28 his later H tones well under the value of H_f , leading to potential confusion of H tones with M
 29 tones. The smaller value of 10.6 Hz avoids this problem, while remaining comfortably above the
 30 threshold value.

31 To summarize, each of the three speakers examined here employs a different strategy for
 32 applying downstep. Only TJ employs the "soft landing" model. FA, in her L-initial sentences,
 33 drops just once and remains at the same level for any subsequent H tone in the span, while TJ
 34 drops twice in two equal steps. We have also seen that absolute values of downstep are adjusted in
 35 the second downstep span in such a way as to maintain the prominence of downstep, while still
 36 keeping H tone values above M tone values. These results, taken together, suggest that what is
 37 important in implementing downstep is not the specific algorithm used but the success of the
 38 algorithm (whatever it may be) in achieving the goal of creating perceptually distinguishable steps
 39 across the downstep domain while avoiding potential mergers with the M tone.
 40

41 ¹⁹In the case of FA, we did not collect data that would allow us to systematically factor out her H tone downdrift
 42 from downstep in later spans.

1 Table 7

Average values of raised (H^+) and unraised sentence-initial H tones in Experiment 4 sentences for speakers TJ and FA

3 Speaker	TJ	FA
5 Raised H (H^+)	111.0	219.5
5 Unraised H (H)	99.5	200.2
7 $\Delta(H^+ - H)$	11.5	19.3
7 $R = H_f$	77.7	170.5
u (Raising factor)	1.53	1.65

9 Values for the raised H^+ are calculated from the average of the first H tone peaks in sentences 7 and 8, and values for the unraised H from the average value of the first H in sentences 4 and 5.

11

11 8.2. *H* tone raising

13

15 We next consider how H raising interacts with downstep. The magnitude of H raising can be estimated by comparing raised vs. unraised H tones in similar contexts. Several of the sentences from Experiment 4 data permit us to do this. Table 7 presents data comparing average values of raised (H^+) and unraised sentence-initial H tones for speakers TJ and FA based on the sentences of Experiment 4. Values for the raised H^+ are calculated from the average of the first H tone peaks in sentences 7 (HLHH) and 8 (HLHM), and values for the unraised H are taken from the average of the first H peaks in sentences 4 (HMHH) and 5 (HMHM).

21 Since we did not record the value of the final H tone (H_f) in these sentences, the value of r shown in the next-to-last row is taken from that of H_f in the H-initial sentences of Experiment 1.

23 How can the H raising factor be defined? A first approach might be to take the H raising increment simply as the difference between H^+ and H , as shown by the line headed $\Delta(H^+ - H)$ in the table. However, it would be convenient, in the absence of evidence to the contrary, to take it as the ratio u (“upstep”) of the respective distances of H^+ and H above the reference line r (that is, $H^+ - r/H - r$), since in that way all the H tones in a downstepping HLHLHL sequence will descend in the same proportion, preserving the downstep ratio. Under this approach, the H raising factor u is calculated as 1.53 for TJ and 1.65 for FA, as shown in the last row of Table 7. The value of a raised H tone (H^+), abstracting from all other factors, is then given by

$$31 \quad H^+ = u(H_b - r) + r, \quad (9)$$

33 where H_b is the base value at which a H tone would theoretically be realized if H raising did not operate on it. (In sentence-initial position, this is just the value of the unraised H as shown in Table 7.)

35 We may now consider how downstep and H raising interact. Let us examine sentences 5 (HMHM), 8 (HLHM), and 9 (HLHL) of Experiment 4, which were earlier graphed in Figs. 10 and 12. In sentence 5, neither factor is expected to apply to either H tone, and both should be realized around their base value H_b . In sentence 8, the first H is raised and the second downstepped. In sentence 9, the first H tone should be raised and the second *both* raised and downstepped. Fig. 12, comparing the f_0 traces of sentences 8 and 9, has already confirmed the latter differences.

41 How can the H tone values in these three sentences be predicted? The value of any H tone is a function of the factors that culminate on it. In a downstepping sequence of H tones, some of which are subject to H raising, the value of the n th H tone (abstracting from all further factors) is given by

43

$$H_n = \langle u \rangle d_{n-1}(H_b - r) + r. \quad (10)$$

The bracketed raising factor $\langle u \rangle$ is selected just in case the target H_n is followed by L. The rest of the expression is adapted from (6), with H_b replacing H_1 , to which it is usually equivalent. We may calculate values of d for Experiment 4 data set on the basis of the values of H_1 and H_2 in sentence 9 and of r in Table 7, giving for example $d = 0.75$ for speaker TJ.

As an illustration, H_2 in TJ's sentence 9 (HLHL), subject to both downstep and raising, is calculated by substituting the appropriate values into (10), giving in this case 102.7 Hz. Table 8 gives the predicted values for sentences 5, 8, and 9 obtained in this way, and compares them with observed values (in parentheses).

The comparisons of interest here are the H_2 values, as well as the H_1 value of sentence 9. We can take the fit to be a good first approximation if any major discrepancies between the predicted and observed values go in the expected direction, given what we know about other interacting factors which have not been modeled here.

In this case, we find two major discrepancies. The first is that the predicted value of H_2 in sentence 8 overestimates the observed value by nearly 4 Hz. This discrepancy can be explained, at least in part, by the lowering effect of M tones in TJ's speech, discussed in connection with Fig. 7 earlier. A further complicating factor here is that Tone Spread (2) causes the tone sequence over these four syllables to be realized as [H–HL–LH–M], in which the second H of the lexical melody HLHM is realized only as the second element of a rising (LH) contour tone; the speaker has little time to reach the expected H target at the end of the third syllable before moving back down to the M tone on the fourth syllable, and may consequently undershoot the H target.²⁰ If this reasoning is correct, the discrepancy here goes in the expected direction. The second major discrepancy involves the two H tones of sentence 9, where the predicted values this time underestimate the expected values by a margin of 4–5 Hz. It will be recalled from the discussion of Experiment 4, however, that TJ is one of the speakers who anticipates the later downstep “at a distance” in this sentence by initializing the first H tones at a higher pitch. This anticipation, persevering into H_2 , would account for most of the discrepancy here.

To summarize, we have seen that the interacting factors of downstep and H raising may target the same H tone, leading to a “compromise” between the values that would be expected if only one or the other of these two factors were at work. Needless to say, the quantitative model proposed here is preliminary and must be confirmed by further experimentation, including explicit modeling of other interacting factors that we have not taken into account here.

8.3. *H resetting*

Let us next consider the H resetting pattern. Table 9 gives a summary of f_0 values for the initial H tone and all following reset H tones (H_r) in the longer sentences (d)–(h) of Experiment 1 data, showing averaged values for L-initial, H-initial and pooled sentences. A H tone is considered reset in these data if it is higher than the preceding H, by however small an amount. (In the very few

²⁰This explanation is supported by a further example of H tone undershoot in exactly the same context in TJ's Fig. 1b data; see Appendix A for discussion.

1 Table 8

Predicted and observed values of the H tones in TJ's sentences 5, 8, and 9 (observed values are parenthesized)

3	Sentence 5 (HMHM)	H_1	99.5 (99.5)
		H_2	99.5 (100.1)
5	Sentence 8 (HLHM)	H_1	111.1 (110.6)
7		H_2	94.1 (90.5)
9	Sentence 9 (HLHL)	H_1	111.1 (115.9)
		H_2	102.7 (106.5)

11 Table 9

Resetting patterns in the mixed-HL sentences of Experiment 1

13	TJ			FA			KG			
	L-in	H-in	pooled	L-in	H-in	pooled	L-in	H-in	pooled	
15	H_1	102.7	100.5	101.6	211.2	223.0	217.1	149.7	146.4	148.0
17	H_{r1}	90.0	90.4	90.2	175.7	181.9	178.8	117.2	115.5	116.4
	H_{r2}	85.1	86.7	85.9	181.0	176.5	178.3	111.9	114.6	113.2
19	H_{r3}	—	83.1	83.1	—	—	—	110.1	111.4	110.8

21 Values are averages in Hz over pooled H- and L-initial sentences (d)–(h). H_1 = first H tone, H_{rn} = n th reset H tone. Only three values are given for FA, who normally reset only twice in our data.

23 cases in which successive H tones rise in pitch, only the final one is counted as reset.) Final tones (H_f) were not included in the averages as they were potentially affected by final lowering.

25 All reset H tones have values well below those of H_1 . As a further trend, each successive H_r is lower than the preceding one in all cases but one (FA's L-initial H_{r2} , resulting from exceptionally high values in the longest sentence in this set). Furthermore, each drop tends to be smaller than that of the preceding one, except for KG's H-initial H_{r3} , based on only one sentence mean. However, these general trends do not hold in all of the individual sentence tokens from which the means were calculated. Also, standard deviations over the ten averaged sentences, though not large (average values TJ = 2.3, KG = 3.0, FA = 5.9), suggest caution in interpreting the very small differences between later pairs of reset values.

33 Keeping these provisions in mind, we note that the resetting pattern of two speakers, KG and TJ, shows some resemblance to the “soft landing” downstep pattern of speaker TJ discussed in the previous section; here it is the successive reset H tones that appear to fall smoothly toward a reference line. To evaluate this interpretation, we may overlay predicted and observed values of reset H tones as we did for downstep in Fig. 13 above. Let us use the symbol r' for the reference line toward which reset H tones appear to fall, and d' for the downstep factor which models the slope of successive reset H tones. For TJ, the value of r' which best fits the data is 83.2 Hz. His downstep factor d' is calculated at 0.38 by expression (11), based on (7) above:

$$41 \quad d' = H_{r1} - r/H_{in} - r. \quad (11)$$

43 For KG, the best fit for r' lies at 112.9 Hz, and d' is calculated at 0.10. It will be noted that the value of r' which best fits the descending H tone values lies well above the final H tone (H_f) values

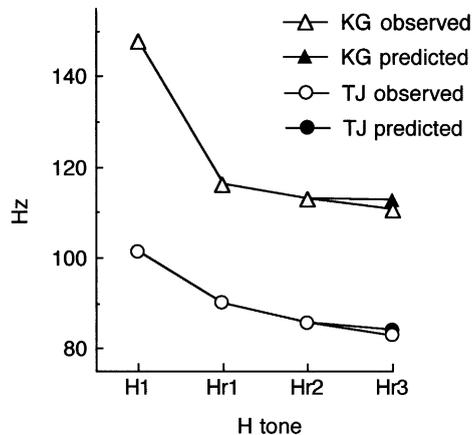


Fig. 14. Plot of observed vs. predicted H tone values in sequences of reset H tones (H_r) for speakers KG and TJ, calculated as for downstep (Fig. 13). Variable settings: for TJ, $r' = 83.17$ Hz, $d' = 0.38$; for KG, $r' = 112.86$ Hz, $d' = 0.10$ Hz.

of both speakers: 77 Hz for TJ, and 83 Hz for KG. Thus, reset H tones will always remain distinct from unreset H tones, no matter how much resetting takes place.

Predicted and observed resetting patterns are shown for KG and TJ in Fig. 14. They show a fairly good overall fit, the main divergence being an overestimation of the third reset H tone H_{r3} by 1–2 Hz for both speakers.

The fact that the divergence goes in the same direction for both speakers suggests, however, that the soft-landing model might not be providing the best possible fit to these data. KG's pattern, and less evidently TJ's, might alternatively be described as involving a single drop to a new H tone ceiling. In this view, the very small subsequent drops (not exceeding 3 or 4 Hz) might be interpreted as resulting from background downdrift operating on H tones, as observed elsewhere. These two interpretations of TJ's and KG's data make nearly identical predictions. However, FA's resetting pattern as shown in Table 9 seems clearly best interpreted in terms of the hard-landing model, since the very small difference between her H_{r1} and H_{r2} values is well accounted for by H tone downdrift, especially prominent in this speaker. That H tone downdrift does operate in the background in FA's speech was already seen in the earlier discussion of FA's mixed-MH sentences (Section 5.2). Her resetting pattern can then be described as a single drop to a value of around 178 Hz for H_{r1} followed by continuation along a gradually downdrifting plateau for any following reset H tone. The value of 178 Hz, like that of the reference line r value estimated for the other speakers, is well above her final H tone value of 169 Hz.²¹

Higher-level downstep patterns such as those observed here, suggesting a fractal-like embedding of smaller patterns into larger patterns of identical structure, have generated some

²¹ Other interpretations of the resetting patterns seem generally less satisfactory. One alternative might claim that the declining values of *all* reset H tones are due to H tone downdrift. However, Fig. 5 has shown that the rate of decline between the first H tone and the first reset H tone is much greater than the rate of decline across H tones in the comparable portions of all-H sentences, at least for the two speakers TJ and FA. The alternative view would have to explain this discrepancy, and would still fail to explain why TJ's resetting patterns fall toward an asymptote instead of showing a linear decline.

1 discussion in the literature. Berg et al. (1992) have observed similar patterns in Dutch, which they
 2 term phrasal downstep, but conclude that they result from a different process than lower-level
 3 downstep (accentual downstep, in their terms). Beckman and Pierrehumbert (1992) have
 4 suggested that such higher-level patterns might be viewed as resulting from the interaction of
 5 various pragmatic factors. We cannot evaluate their proposal on the basis of our present
 6 experimental corpus, which was not designed to explore such factors, but it calls for an
 7 examination of the effect of various pragmatic and discourse-level factors on f0 patterns in future
 8 work on Yoruba.

9 A related issue, also bearing on the interpretation of higher-level downstep, involves variability.
 10 Our modelling has been based on mean sentence values rather than individual tokens, and we
 11 have not carried out a statistical evaluation of token-to-token variability. Even in our averaged
 12 data, we find considerable variability among the sentences on which the means were based. Thus,
 13 the higher-level downstep patterns represent statistical trends that are not necessarily apparent in
 14 any given sentence. Even with these qualifications, it seems that some form of higher-level
 15 downstep characterizes the productions of all of our speakers, suggesting that we may be dealing
 16 with a single principle, implemented in different ways.

17 To summarize, just as in the case of downstep, we have found somewhat different patterns of f0
 18 decline across reset H tones from one speaker to another. In spite of these differences, the specific
 19 quantitative choices made by each one, like those for the variables *d* and *r* discussed earlier,
 20 appear to have a functional significance: they guarantee that no matter how long a sentence may
 21 be, nor how many times a speaker resets, there will always be room for further resetting and
 22 downstepping within the intonational unit, and that H tones will always remain distinct from M
 23 tones. In this sense, resetting and downstep “conspire” to guarantee the optimal utilization of the
 24 speaker’s pitch space, making any additional look-ahead strategy unnecessary.

25 8.4. *L assimilation*

26
 27 One of the more surprising results of Experiment 1 data concerned the realization of L tones.
 28 For three speakers (TJ, KG, FA), the initial L tones of L-initial sentences were pitched very low
 29 (Figs. 2–4). For two of these (TJ, KG), the first L tone following a H tone was considerably raised
 30 with respect to this initial tone, and the one or two L tones following it showed declining values
 31 (Figs. 2–3). For speakers TJ and FA, no later L tones showed any decline greater than that which
 32 could be attributed to L tone downdrift (Fig. 5). The controlled comparisons of Experiment 4
 33 confirmed the absence of downstep on L tones, for all four speakers (Table 5).

34 All these observations converge to suggest that L tones do not undergo downstep, but “carry-
 35 over” assimilation from a preceding H tone in the sense of Gandour et al. (1994). In this account,
 36 the fact that a sentence-initial L tone is low-pitched for all speakers is explained by the absence of
 37 a preceding H-tone assimilation trigger. That the first few L tones following H tones are raised for
 38 two speakers is due to carry-over assimilation from the H’s. That these L tones show a declining
 39 pattern follows from the fact that the H tone assimilation triggers themselves show a declining
 40 pattern, if we may assume that assimilation is proportional to the height of the trigger. The
 41 evidence strongly suggests, therefore, that L tones are subject to a local factor of carry-over
 42 assimilation conditioned by a preceding H tone, but not to downstep. Any remaining tendency for
 43 L tones to decline can be attributed to L tone downdrift operating in the background.

1 L assimilation is subject to conditions varying somewhat from speaker to speaker; thus TJ, for
 2 instance, appears to show assimilation only to raised H tones. As it is not central to the subject of
 3 this study, we do not attempt quantitative modeling of L assimilation here.

5 8.5. *Why downstep and H raising?*

7 At the outset of this section, we raised the question: why should downstep and H raising coexist
 9 in a tone language like Yoruba? Or alternatively, why should one exist, given the other? We shall
 11 suggest that the answer has much to do with the fact that Yoruba is a three-level tone language,
 13 with phonologically distinctive H, M, and L tone levels. If only H raising existed without
 15 downstep, H tones would be pitched unnecessarily high throughout an utterance, serving no
 17 detectable function; considerations of articulatory economy would suggest that such a situation, if
 19 it ever existed historically, would be unstable and unlikely to survive for long. Downstep has the
 21 effect of lowering noninitial, raised H tones to values where they may be comfortably produced. If
 on the other hand, downstep existed without H raising, downstepping H tones would soon merge
 with M tones in longer utterances (if M were downstepped as well, these would tend to merge with
 L tones). H raising has the advantage of keeping downstepping H tones from crossing into the
 frequency band reserved for M tones, allowing distinctive H vs. M contrasts to be maintained
 throughout the utterance. Similarly, resetting typically applies at points at which downstepped H
 values have fallen under the expected value of nondownstepped H tones, and so it, too,
 contributes to keeping H tones out of the M tone frequency band.

23 By way of illustration, Fig. 15 overlays the mixed-HL and all-M sentences (c) from Experiments
 25 2 and 3 for TJ and FA. It shows that for both speakers, downstepping H tones do not cross into
 27 the M tone band. TJ keeps his downstepping H tones well above the M tone level, while FA
 allows her H tones to approach it, without crossing into it. The role of resetting in keeping these
 tones apart is especially apparent in the comparison between FA's third and fourth H tones.

29 It may appear that FA's H tones are getting perilously close to the M tones in the second half of
 31 her sentences. However, other factors help to keep them apart. Further data in Experiment 4 show
 33 that in FA's speech, downstep affects not only a H tone but also an M tone following it. Thus, for
 35 example, the average value of the M tone following the downstepped H tone in FA's sentence 8
 (HLHM) is about 7 Hz lower than that of the comparable M tone following the nondownstepped
 H tones in her sentences 2 (HHHM) and 3 (HMHM).²² This extra lowering of M tones after
 downstepped H tones helps to create a buffer between H and M tones in longer downstep spans.
 Recall also that for all speakers, H tones are realized as rising tones after L tones due to the
 operation of Tone Spread (2), while M tones, unaffected by Tone Spread, remain level. All these
 factors combine to help keep H tones distinct from M tones after L tones.

37 From a typological point of view, we conclude that although Yoruba is correctly described as a
 39 "terraced-level tone language" insofar as it possesses downstep, it is also a "discrete-level tone
 41 language" in the sense that it keeps its tone levels phonetically distinct. (For these concepts, see
 Welmers 1973, pp. 81–83.) Thus, as suggested earlier by Connell and Ladd (1990), these two
 typological categories are not mutually exclusive.

43 ²²In contrast, TJ, whose H and M tones are well separated in Fig. 15, shows only a very small lowering of M tones
 (1–3 Hz) in the same context.

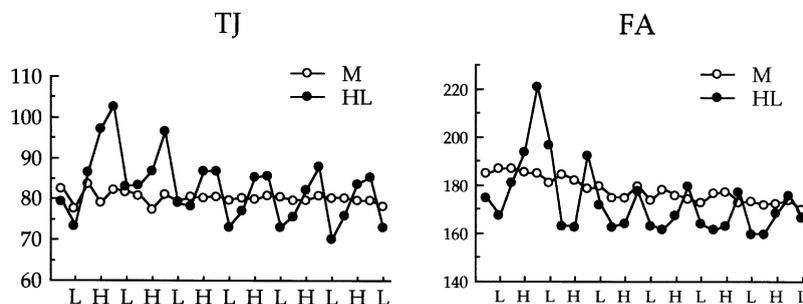


Fig. 15. Superimposed graphs of the longest mixed-HL and all-M sentences of Experiments 2 and 3 for TJ and FA. It can be observed that downstepping H tones do not descend below the M line (values in Hz).

9. Conclusions

This study has motivated a compositional approach to tone production in which the overall shape of an f_0 contour is viewed as resulting not from any single factor (such as downstep) but from the interaction of a number of factors, each of which can vary independently of the others to at least some degree. Earlier approaches to the study of pitch and intonation based on a similar view include Thorsen (1983), Vaissière (1983), Pierrehumbert and Beckman (1988), Clements (1990), Berg et al. (1992), and Ladd (1996). The major new contribution of our study has been to confirm the interaction of H raising and other factors with downstep in accounting for f_0 contours in a tone language, Yoruba. Furthermore, by extracting H raising as an independent effect, we have been able to isolate the properties of downstep as such and show that for some speakers, it is similar to downstep as it has been described in English, Japanese, Dutch, and other nontonal languages. This result supports the widely assumed, but previously unsubstantiated view that downstep in nontonal languages may be of the same fundamental nature as downstep in African languages.

Other effects besides downstep have been found to contribute to f_0 decline within and across the utterance. These include “background” downdrift across single tone levels, long-distance anticipatory raising of initial H tones before LHL sequences, and local “carry-over” assimilation. H resetting resembles downstep in that it resets the ceiling for following H tones, though its implementation is quantitatively different from downstep. Another result of this study is that speakers vary considerably in their implementation of downstep, some using a model of gradual decay toward a reference line while others drop in a single step, or two steps of equal magnitude. In spite of these differences, these various approaches interact with resetting to produce f_0 contours in which there is always room for further, perceptually prominent downsteps, while insuring that H tones will not penetrate the frequency band reserved for M tones.

While we have found evidence that tone production in tone languages makes reference to abstract “tone levels” such as H and L, as has been proposed in earlier work (Clements, 1979), we have found no evidence for the manipulation of the full set of tone levels as a whole, corresponding to what some linguists have called “register shift” or “key lowering”—i.e., a shift of all tone levels downward or upward as a block. That this does not take place in Yoruba is shown by the fact that downstep of H tones does not entail downstep of L tones (see Table 5). It is still an open question, however, whether downstep of H tones entails downstep of following M tones (but see discussion of data from FA following Fig. 15). The application of downstep to both H and M registers at once, if this can be

1 confirmed for other speakers, would provide a further way of keeping these tones apart, and would
 2 tend to credit the idea that downstep involves range contraction, by lowering the ceiling (but not the
 3 floor) of the frequency band within which tone levels are distributed. This question should be studied
 4 systematically in the future. Finally, we have found no strong evidence for any global, utterance-level
 5 declination in Yoruba, beyond the effects of downstep and what we have termed “background
 6 downdrift” across individual tone levels.

7 Further questions that deserve attention in future work on Yoruba include the following:

- 9 ● To what extent does the present account, based on a small number of selected tone sequences,
 10 carry-over to sequences not examined here, including those containing M tones, contour tones, and
 11 longer sequences of consecutive H tones? What is the special behavior of tones at boundaries?
- 12 ● What is the effect of syntactic structure on tone production? In particular, to what extent do
 13 resetting points coincide with syntactic (word, phrase, clause) boundaries? Within what
 14 prosodic domains do downstep and downdrift apply?
- 15 ● To what extent do f_0 patterns vary with sentence type (statements, questions, vocatives, exhortations,
 16 commands, etc.), information content (topic vs. comment, location of focus, parentheticals, etc.),
 17 speech styles (formal vs. informal, read vs. improvised), and different speech rates?
- 18 ● What is the effect of systematically varying the amplitude or pitch range of a given utterance?
 19 To what extent do variations in f_0 convey different illocutionary and expressive functions?

21 Our discussion has also not addressed the question of whether H raising and downstep are
 22 purely phonetic principles in Yoruba, or whether in spite of their largely nondistinctive nature
 23 they have become phonologized, creating phonologically raised and downstepped tones, as has
 24 happened in the course of evolution of many other African languages (see [Snider \(1998\)](#) for
 25 pertinent discussion). Going beyond Yoruba, we would like to know how downstep, H raising
 26 and other principles interact in other languages with three tone levels or more. All these issues
 27 remain open, showing that the study of African tone languages continues to have rich potential
 28 for our understanding of speech prosody.

31 Acknowledgements

33 This study draws upon experimental results first presented in Laniran’s Ph.D. thesis (1992),
 34 augmented with new data as discussed in Section 2.2. Though listed in antialphabetical order,
 35 both authors contributed equally to the elaboration of the present article. Earlier versions were
 36 presented at the 23rd Annual Conference on African Linguistics, Ohio State University,
 37 Columbus, Ohio, the 28th Colloquium on African Languages and Linguistics, University of
 38 Leiden, and at colloquia at the University of Edinburgh and the University of Paris 3 (Sorbonne-
 39 Nouvelle). We thank participants at those occasions, in particular Akin Akinlabi, Mary Beckman,
 40 Chet Creider, Bob Ladd, Ian Maddieson, and Annie Rialland, for useful questions and
 41 suggestions. Special thanks are due to Mary Beckman, Bruce Connell, and Nina Grønnum (the
 42 first two in their capacity of nonanonymous reviewers) for their thoughtful and constructive
 43 comments on earlier drafts, which have helped us avoid certain errors and make many material
 improvements. Any remaining errors are of course our own.

1 A second set of “H-initial sentences” sentences, identical to those above except that the L-initial
 2 name *Ògúnbòdé* was replaced by the H-initial name *Lábòdé*, was added to control for the possible
 3 downstepping influence of a sentence-initial L tone on following tones. In the Yoruba
 4 orthography used here, vowel letters followed by *n* designate phonemically nasal vowels and
 5 the subscripted vowels *ɛ ɔ* represent the lower mid vowels [ɛ ɔ].

6 All sentences have identical syntactic structure in which the proper noun subject is followed by
 7 a predicate beginning with a contracted verb + noun sequence. *Gbà gbá* is the contraction of the
 8 verb /gbà/ ‘receive’ and the noun /gbá/ ‘eggplant’, created through a regular process of vowel
 9 elision. In the (b) sentences in both sets, the noun *gbáa* bears an M tone as the second component
 10 of a H-to-M falling tone. This M tone, originating in a floating M tone occurring at the juncture
 11 between the first and second noun of N+N genitival constructions (Akinlabi, 1985), is not
 12 indicated in the orthography, and was first noticed only after the recordings were made. Its
 13 presence between the surrounding H and L tones blocks the process of H raising that applies only
 14 to H+L sequences, as explained in the text. In all other sentences, *gbá* bears a simple H tone.

15 In the L-initial test sentences (a) and (c-h) given to FA, the noun *Bàbágbèmi* was used in place
 16 of *Ògúnbòdé*. A new (b)-sentence, *Bàbá/Wón gbà gbá bọ̀ fún Bòdé àbá*/they received an eggplant
 17 and brought it back for Bode’, was created for FA in order to eliminate the M tone of *gbáa*.

19

Appendix B. Test sentences used in Experiment 2

21

22 Experiment 2 involved three sets of sentences, all-H, all-L, and mixed-HL. Each set contained
 23 three sentences, ranging from short (a) to long (c). Tones belonging to the parenthesized parts of
 24 the sentence frame were not included in the graphs. All sentences are listed below in the form in
 25 which they were presented to TJ. Sentence (2a) was replaced with a longer sentence for FA so that
 26 all her set (a) sentences were of equal length.

27

1. All-H sentences:

29

a. Gbòláválé gbé Bádé (àṭọ́lá lálé.)

‘Gbòlawale carried Bádé (and Ọ́lá at night)’

31

b. Gbòláválé gbé Bọ́lá Bádé (àṭọ́lá lálé)

‘Gbòlawale carried Bọ́lá, Bádé (and Ọ́lá at night)’

33

c. Gbòláválé gbé Bádé, Bọ́lá, Wálé, Gbádé (àṭọ́lá lálé)

‘Gbòlawale carried Bádé, Bọ́lá, Wale, Gbade (and Ọ́lá at night)’

35

2 All-L sentences:

37

a. Dàpọ̀ kò rẹ̀bà (àtedé lálé)

‘Dapọ̀ did not buy *ẹ̀ba* (and shrimp at night)’

39

b. Dàpọ̀ kò rọ̀gẹ̀dẹ̀ gbẹ̀gìrì (àtedé lálé)

NEG buy-bananas

‘Dapọ̀ did not buy bananas, *gbẹ̀gìrì* (and shrimp at night)’

41

c. Dàpọ̀ kò rọ̀gẹ̀dẹ̀, gbẹ̀gìrì, dòdò, ẹ̀bà (àtedé lálé)

‘Dapọ̀ did not buy bananas, *gbẹ̀gìrì*, plantains, *ẹ̀ba* (and shrimp at night)’

43

3. Mixed HL sentences

a. Ògúnbòdé gbà gbá à (tedé lálé)

- 1 ‘Ogunbode received an eggplant (and shrimp at night)’
 2 b. Ògúnbòdé gbà, gàrí à (tedé lálé)
 3 ‘Ogunbode received an eggplant, *gari* (and shrimp at night)’
 4 c. Ògúnbòdé gbà gbá, è. g é., àdá, gàrí à (tedé lálé)
 5 ‘Ogunbode received an eggplant, cassava, *gari* (and shrimp at night)’
 6

9 Appendix C. Test sentences used in Experiment 3

11 Experiment 3 involved sentences containing M tones, as shown below. As before, each set
 12 contained three sentences ranging from short (a) to long (c). These sentences are listed below in
 13 the form in which they were presented to TJ. Sentence (1a) was replaced with a longer sentence for
 14 FA so that all her set (a) sentences were of equal length.
 15

- 16 1. All-M sentences:
 17 a. Dàpò kò rògèdè gbègìrì (àtedé lálé)
 18 NEG buy-bananas
 19 ‘I hawked fish (and shrimp at night)’
 20 b. Mo kiri eja (àtedé lálé)
 21 I hawk fish and-shrimp at-night
 22 ‘I hawked fish, meat, palm-wine (and shrimp at night)’
 23 c. Mo kiri eja, eran, emu, eyin, oyin (àtedélálé)
 24 ‘I hawked fish, meat, palm-wine, eggs, honey (and shrimp at night)’
 25 2. MH sentences:
 26 a. Oláyemíbe Gbémi (àtọlá lálé)
 27 ‘Olayemi appealed to Gbemi (and Ola at night)’
 28 b. Oláyemíbe Gbémi, Lánre (àtọlá lálé)
 29 ‘Olayemi appealed to Gbemi, Lanre (and Ola at night)’
 30 c. Oláyemíbe Gbémi, Mọyẹ, Lánre, Fẹmi (àtọlá lálé)
 31 ‘Olayemi appealed to Gbemi, Mọyẹ, Lanre, Fẹmi (and Ola at night)’
 32 3. LM sentences:
 33 a. Mo kiri eja, eran, emu (àtedé lálé)
 34 meat palm-wine
 35 ‘I bought a knife, medicine (and shrimp at night)’
 36 b. Mo ròbẹ, àga, àgbo, óbọ (àtedé lálé)
 37 ‘I bought a knife, a chair, medicine, a monkey (and shrimp at night)’
 38 c. Mo ròbẹ, àwo, àga, ibọn, àgbo, óbọ (àtedé lálé)
 39 ‘I bought a knife, a plate, a chair, a gun, medicine, a monkey (and shrimp at night)’
 40 4. LMH sentences:
 41 a. Gbọlá bi Bùnmi, Gbémi àtọ(lá lálé)
 42 ‘Gbọla asked Bunmi, Gbemi (and Ola at night)’
 43 b. Gbọlá bi Bùnmi, Gbémi, Dẹjọ, Lánre àtọ(lá lálé)
 44 ‘Gbọla asked Bunmi, Gbemi, Dẹjọ, Lanre (and Ola at night)’

1

3

Appendix D. Test sentences used in Experiment 4

5

All sentences in this set have identical syntactic structure, in which a proper-noun subject is followed by a serial verb predicate introduced by a contracted V+N sequence (e.g. *relá* is the contraction of /ré/ ‘pick’ + /ilá/ ‘okra’, created through a regular process of vowel elision). Glosses in each set are ‘X (picked okra/bought mushrooms/bought books) for Láyemí,’ respectively. The first five syllables contain only sonorants, to eliminate the perturbing effects of obstruents. All possible instantiations of the schema HXHY followed by a constant H tone are represented in the data. All sentences were produced in the frame *bò fún Láyemí* ‘for Láyemí’ to absorb final lowering effects; the tone pattern of this frame, LHHMH, is not plotted on the graphs.

15 Set A: HHHY

1. HHHH: Déwálé relá

17 2. HHHM: Déwálé rolú

3. HHHL: Déwálé ràwé

19 Set B: HMHY

4. HMHH: Móyemí relá

21 5. HMHM: Móyemí rolú

6. HMHL: Móyemí ràwé

23 Set C: HLHY

7. HLHH: Máyómí relá

25 8. HLHM: Máyómí rolú

9. HLHL: Máyómí ràwé

27

29

References

31

- Akinlabi, A. (1985). Underspecification of tones in Yoruba. *Ph.D. thesis*, University of Ibadan.
- 33 Akinlabi, A., & Liberman, M. (1995). On the phonetic interpretation of the Yoruba tonal system. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the 13th international congress of phonetic sciences*, Vol. 1, Stockholm (pp. 42–45).
- 35 Bamgbose, A. (1966). The assimilated low tone in Yoruba. *Lingua*, 16, 1–13.
- Beckman, M., & Pierrehumbert, J. (1992). Comments on chapters 14 and 15. In G. J. Docherty, & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 387–397). Cambridge: Cambridge University Press.
- 37 Berg, R. van den, Gussenhoven, C., & Rietveld, T. (1992). Downstep in Dutch: Implications for a model. In G. J. Docherty, & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 335–359). Cambridge: Cambridge University Press.
- 39 Clements, G. N. (1979). The description of terraced-level tone languages. *Language*, 55(3), 536–558.
- 41 Clements, G. N. (1990). The status of register in intonation theory: Comments on the papers by Ladd and by Inkelas and Leben. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology 1: Between the grammar and the physics of speech* (pp. 58–71). Cambridge: Cambridge University Press.
- 43 Clements, G. N., & Ford, K. C. (1979). Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry*, 10(2), 179–210.

- 1 Connell, B., & Ladd, D. R. (1990). Aspects of pitch realization in Yoruba. *Phonology*, 7, 1–30.
- 2 Courtenay, K. (1971). Yoruba: A 'terrace-level' language with three tonemes. *Studies in African Linguistics*, 2(3), 239–
- 3 255.
- 4 Farnetani, E. (1997). Coarticulation and connected speech processes. In W. J. Hardcastle, & J. Laver (Eds.), *The*
- 5 *handbook of phonetic sciences* (pp. 371–404). Oxford: Blackwell.
- 6 Gandour, J., Potisuk, S., & Dechongkit, S. (1994). Tonal coarticulation in Thai. *Journal of Phonetics*, 22, 477–492.
- 7 Hombert, J.-M. (1974). Universals of downdrift: Their phonetic basis and significance for a theory of tone. *Studies in*
- 8 *African Linguistics*, (Suppl. 5), 169–183.
- 9 Hombert, J.-M. (1977). Consonant types, vowel height, and tone in Yoruba. *Studies in African Linguistics*, 8, 173–190.
- 10 Hyman, L. (1993). Register tones and tonal geometry. In H. van der Hulst, & K. Snider (Eds.), *The phonology of tone:*
- 11 *The representation of tonal register* (pp. 75–108). Berlin: Mouton de Gruyter.
- 12 Inkelas, S., & Leben, W. R. (1990). Where phonology and phonetics intersect: The case of Hausa intonation. In J.
- 13 Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology 1: Between the grammar and the physics of speech*
- 14 (pp. 17–35). Cambridge: Cambridge University Press.
- 15 Ladd, D. R. (1984). Declination: A review and some hypotheses. *Phonology Yearbook*, 1, 53–74.
- 16 Ladd, D. R. (1990). Metrical representation of pitch register. In J. Kingston, & M. Beckman (Eds.), *Papers in*
- 17 *laboratory phonology 1: Between the grammar and the physics of speech* (pp. 35–57). Cambridge: Cambridge
- 18 University Press.
- 19 Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- 20 Laniran, Y. O. (1992). Intonation in tone languages: The phonetic implementation of tones in Yoruba. *Ph.D. thesis*,
- 21 Cornell University.
- 22 Laniran, Y. O. (1995). Implementing a floating tone. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the 13th*
- 23 *international congress of phonetic sciences*, Vol. 2 (pp. 390–393). Stockholm: KTH and Department of Linguistics,
- 24 Stockholm University.
- 25 Laniran, Y. O., & Clements, G. N. (1995). A long-distance dependency in Yoruba tone realization. In K. Elenius, & P.
- 26 Branderud (Eds.), *Proceedings of the 13th international congress of phonetic sciences*, Vol. 2 (pp. 734–737).
- 27 Stockholm: KTH and Department of Linguistics, Stockholm University.
- 28 Laniran, Y. O., & Gerfen, C. (1997). High raising, downstep and downdrift in Igbo. *Paper presented at the 71st annual*
- 29 *meeting of the Linguistic Society of America*, Chicago.
- 30 La Velle, C. R. (1974). An experimental study of Yoruba tone. *Studies in African Linguistics*, (Suppl. 5), 185–194.
- 31 Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M.
- 32 Aronoff, & R. T. Oehrle (Eds.), *Language and sound structure* (pp. 157–233). Cambridge, MA: MIT Press.
- 33 Liberman, M., Schultz, J. M., Hong, S., & Okeke, V. (1993). The phonetic interpretation of tone in Igbo. *Phonetica*, 50,
- 34 147–160.
- 35 Lindau, M. (1986). Testing a model of intonation in a tone language. *JASA*, 80(3), 757–764.
- 36 Maeda, S. (1976). A characterization of American English intonation. *Ph.D. dissertation*, MIT, Cambridge, MA.
- 37 Mountford, K. (1983). Bambara declarative sentence intonation. *Research in phonetics*. Vol. 3. Bloomington: Indiana
- 38 University.
- 39 Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. *Ph.D. thesis*, MIT, Cambridge, MA.
- 40 Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. *Linguistic Inquiry Monograph 15*. Cambridge, MA:
- 41 MIT Press.
- 42 Pike, K. L. (1948). *Tone languages*. Ann Arbor: the University of Michigan Press.
- 43 Pike, K. L. (1966). *Tagmemic and matrix linguistics applied to selected African languages*. Washington, DC: Department
- of Health, Education, and Welfare (Office of Education, Bureau of Research).
- Pulleyblank, D. (1986). *Tone in lexical phonology*. Dordrecht: Reidel.
- Rabiner, L. R., & Schafer, R. W. (1975). *Digital processing of speech signals*. Englewood Cliffs, NJ: Prentice-Hall.
- Rialland, A. (1981). Le système tonal du gurma (langue gur de Haute-Volta). *Journal of African Languages and*
- 44 *Linguistics*, 3, 39–64.
- 45 Rialland, A. (1983). Le système tonal du moba comparé à celui du gurma. In J. Kaye, H. Koopman, D. Sportiche, & A.
- Dugas (Eds.), *Current approaches to African linguistics*, Vol. 2 (pp. 217–234). Dordrecht: Foris.

- 1 Rialland, A. (1997). Le parcours du “downstep” ou l’évolution d’une notion. *Bulletin de la Société de Linguistique de*
Paris, 92.1, 207–243.
- 3 Rialland, A. (2001). Anticipatory raising in downstep realization: Evidence for preplanning in tone production. In S.
 Kaji (Ed.), *Proceedings of the symposium on cross-linguistic studies of tonal phenomena: Tonogenesis, Japanese*
accentology, and other topics (pp. 301–321). Tokyo: Institute for the Study of Languages and Cultures of Asia and
 5 Africa (ILCAA), Tokyo Institute of Foreign Studies.
- 7 Rialland, A., & Somé, P. A. (2000). Dagara downstep: How speakers get started. In V. Carstens, & F. Parkinson (Eds.),
Advances in African linguistics. Trends in African Linguistics, Vol. 4 (pp. 251–263). Trenton, NJ: Africa World Press.
- 9 Schachter, P. (1965). Some comments on J.M. Stewart’s “The typology of the Twi tone system.” *Bulletin of the Institute*
of African Studies, Vol. 1, Legon, Ghana, unpublished (preprint).
- 11 Snider, K. L. (1998). Phonetic realization of downstep in Bimoba. *Phonology*, 15(1), 77–101.
- 13 Stewart, J. M. (1965). The typology of the Twi tone system. *Bulletin of the Institute of African Studies*, Vol. 1, Legon,
 Ghana, unpublished (preprint).
- 15 Stewart, J. M. (1993). Dschang and Ebrié as Akan-type total downstep languages. In H. van der Hulst, & K. Snider
 (Eds.), *The phonology of tone: The representation of tonal register* (pp. 185–244). Berlin: Mouton de Gruyter.
- 17 Thorsen, N. (1983). Standard Danish sentence intonation: Phonetic data and their representation. *Folia Linguistica*, 17,
 187–220.
- 19 Vaissière, J. (1983). Language-independent prosodic features. In A. Cutler, & D. R. Ladd (Eds.), *Prosody: Models and*
Measurements (pp. 53–66). Berlin: Springer.
- 21 Ward, I. C. (1952). *An introduction to the Yoruba language*. Cambridge: Cambridge University Press.
- Welmers, W. E. (1973). *African language structures*. Berkeley: University of California Press.
- Xu, Y. (1997). Contextual tonal variations in mandarin. *Journal of Phonetics*, 25, 61–83.