Journal of Phonetics (2002) **30**, 101–129
doi:10.1006/jpho.2001.0156
Available online at http://www.idealibrary.com on IDEAL®

Tone languages and the universality of intrinsic *F*0: evidence from Africa

Bruce Connell

Department of Languages, Literatures, and Linguistics, York University, 4700 Keele Street, Toronto, ON, Canada M3J 1P3

Received 28th July 2000, accepted 9th November 2001

A correlation between vowel height and fundamental frequency, whereby high vowels have higher *F*0 than low vowels, is said to be universal. The available evidence suggests that this intrinsic *F*0 (IF0) extends even to tone languages, which might be expected to control or constrain *F*0. Little work, however, has been done on IF0 in tone languages of Africa, where the tone systems are in certain respects more complex than those found elsewhere. This paper presents new research on four African tone languages which permits several questions concerning IF0 in these languages to be addressed. The central question, whether tone inventory size is a constraining factor on IF0, presupposes that IF0 does indeed exist in tone languages. Reports in the literature that IF0 is reduced or neutralized for Low tone suggest that there may be an IF0 gradient with respect to tone height. Results confirm the existence of IF0 for three of the four languages studied, but also suggest that it may be constrained in some tone languages. Tone inventory size alone, however, does not account for this; rather, the nature of the tone system, and in particular the degree of *F*0 modulation used in producing tonal contrasts appears to the primary factor. While IF0 is generally reduced for Low tone, not all of these four languages show the postulated gradient. This finding fits with research suggesting that a different physiological mechanism may be associated with the production of low range *F*0.

1. Introduction

A correlation between vowel height and fundamental frequency is well known, the essential claim being that high vowels have higher intrinsic fundamental frequency (IF0) than low vowels. Whalen & Levitt (1995) summarize the available literature, and elaborate this basic finding to include the following claims:

- There is no significant difference in IF0 between front and back vowels. This follows from the assumption that IF0 is a function of vowel height. Whalen and Levitt note a tendency for [u] to show higher *F*0 than [i]; Reinhold Petersen (1978) suggests that this is a result of active lowering of the larynx in producing [u].

E-mail: bconnell@yorku.ca

0095-4470/02/010101 + 29 $35.00/0

- IF0 is gradient; e.g., in a vowel system with three degrees of height, mid vowels will have an IF0 between that of high and low vowels. While Whalen and Levitt (p. 355) suggest this to be the case, it is not actually tested in their work, which was restricted to looking only at high and low vowels. However, assuming that IF0 is tied to vowel height, and that perceived vowel height correlates with tongue height, a gradient is to be expected. The detailed analysis of IF0 by Fischer-Jorgensen (1990) for six German speakers shows this to be the case, but only when tense and lax vowels are considered separately. Otherwise, the correlation breaks down, with factors such as jaw height also apparently playing a role. The results by Iivonen (1989) for German largely confirm those of Fischer-Jorgensen.

The survey by Whalen and Levitt includes a number of tone languages, which also appear to be subject to IF0. These allow the following additional observations:

- IF0 is neutralized for Low tones, i.e., differences in *F*0 between high and low vowels are reported for vowels with High tone, as found for nontonal languages, but no (or relatively small) differences were found for vowels with Low tones.[1] This finding agrees with the speculation by Ladd and Silverman (1984) that low pitch accounts for certain cases where they found no IF0. It is, however, not without exception. Whalen and Levitt (p. 358) mention Kammu as the one exception in the languages they surveyed, but at least one other tone language which they include, Iṣẹkiri, also appears to buck the general trend, as discussed below.
- IF0 is gradient with respect to tone; i.e., in languages with more than two tone levels, IF0 should decrease as tone level moves from High to Mid to Low. Although there is little data available in the literature to support or refute this claim, it follows from the discussion of Kammu by Whalen and Levitt, that a mid-range pitch may be used in the realization of Kammu L, accounting for its failure to show a neutralizing influence.

The first of these two claims, that IF0 is neutralized for Low tones, seems relatively well substantiated by the studies reviewed in Whalen and Levitt; the second, together with the earlier mentioned vowel-gradient effect, is less well substantiated since many of the studies of tone languages did not include an appropriate range of tones or vowels to permit examination of these questions.

Whalen and Levitt offer no discussion of the possible influence of size of tone inventory, perhaps because they did not have enough tone languages of similar type and a sufficient range of inventories to draw any conclusions. It is reasonable to suggest, however, that this may indeed be a factor which affects IF0: whereas with vowel inventory, size was considered a possible influence with larger inventories more prone to utilize IF0 as a means of enhancing vowel contrasts, larger tone inventories might constrain IF0 so as not to endanger tonal contrasts.

Considerable attention is paid by Whalen and his colleagues (Whalen & Levitt, 1995; Whalen, Levitt, Hsiao & Smorodinsky, 1995; Whalen, Gick, Kumada & Honda, 1998) to the question of whether IF0 is an enhancement effect, and therefore phonological in nature, or whether it is a universal phonetic effect. The main proponents of the enhancement view are Diehl and Kingston (Diehl & Kluender, 1989; Diehl, 1991; Kingston, 1993; Kingston & Diehl, 1994), who argue that IF0 is utilized to increase the salience of vowel

[1]Throughout the paper tone labels are written using upper case initials (High = H, Mid = M, Low = L, etc.) to distinguish them from vowel heights, high, mid, and low.

contrasts. Whalen and his colleagues provide well-reasoned counter-arguments to this claim, which need not be rehearsed here. It is worth adding, though, that were IF0 to be used to enhance vowel contrastivity in a tone language, this could only be done by potentially putting tonal contrasts at risk.

Accepting that IF0 is a universal phonetic effect, however, does not rule out the possibility that it may also be subject to constraints. That phonological considerations may play a role in constraining, or possibly exploiting, a process is one language which appears to be automatic in others seems well supported. With respect to *F*0 it is, for example, generally accepted that the perturbation associated with the release of a consonant is a universal phonetic effect (e.g., Ohde (1984); cf. Kingston and Diehl (1994), who offer a contrasting view). However, it appears that this may be constrained in some tone languages. Gandour (1974) and Hombert (1977, 1978) report similar durations of *F*0 perturbations for two tone languages, Thai and Yoruba, respectively, of about 50 ms. This is considerably shorter than the durations reported for English by Hombert, where significant differences between voiced and voiceless consonants were found to exist even after 100 ms. It has also been claimed that *F*0 perturbations may be exploited to enhance phonological contrasts (Kingston & Diehl, 1994). Jun (1993, 1996) shows segmental effects of *F*0 to be greater in Korean, in both magnitude and duration, than in languages such as French and English, and argues that the effect is phonologized to enhance the three-way Korean consonant contrast in voicing. Similarly Jang (2000), again with respect to Korean, reports the nature of segmental effects on *F*0 to be significantly different depending on consonant type, and considers his results to be a positive indication for the use of segmental *F*0 in consonant identification.

This is also evidence that other phonetic aspects of pitch realization may be subject to phonological influence. Declination, the automatic lowering of *F*0 over the course of an utterance (Pierrehumbert, 1980; Ladd, 1984), is widely agreed to be a universal phonetic effect (and perhaps not even species specific; Hauser & Fowler, 1992), but it is suspended in some languages where the robustness of tonal contrasts appears threatened. Hombert (1974) establishes an implicational hierarchy for register tone languages, such that declination (subsumed by his "downdrift") only occurs with higher tones if it is also present for lower tones (see also Connell, 1999*a*, *b*). This is not only a question of declination being suspended in interrogatives as opposed to declaratives as reported, for example, by Lindau (1986) and Inkelas and Leben (1990) for Hausa (see also Ladd, 1996), and therefore it is not attributable to a "marked" *vs.* "unmarked" opposition (cf. Bolinger, 1964). The phenomenon described by Hombert applies to sentences of the same type (e.g., declaratives), where there is no basis for suggesting that Highs are marked relative to Mids which are marked relative to Lows.

Specifically with respect to *F*0 differences across vowels, research shows that IF0 is constrained in singing as compared to speech (Ternström, Sundberg & Colldén, 1988; Fowler & Brown, 1997; Hollien, Mendes-Schwartz & Nielsen, 2000). While there is no agreement as to the mechanism responsible for this reduction (see Sapir (1989) for discussion), there does appear to be consensus that "singing requires a high level of precision of *F*0 control" (Sapir, 1989, p. 48). Greater control of *F*0 may equally be required in tone languages as opposed to nontone languages, and quite conceivably more in some than others, depending on the nature of the tone system. It is reasonable to suggest, then, that IF0 may also be constrained in tone languages, particularly in contexts where a degree of variation such as is reported elsewhere might jeopardize tonal contrasts. In other words, in languages with a more "crowded" tonal space, IF0 might be constrained, or more tightly controlled. Investigating languages with different sizes of tone inventory would provide insight into this question.

In summary, even in accepting the universal phonetic nature of IF0, with respect to tone languages there are several issues of interest that require further examination or clarification:

- whether the size of tone inventory has an influence on IF0,
- whether, or to what extent, Low tone neutralizes IF0,
- whether there is an IF0 gradient by tone.

In this paper, we address the issue of IF0 in tone languages, drawing on new evidence from four African languages, as well as information available on IF0 in other African languages. Our primary focus is on testing as to whether IF0 is indeed necessarily present in tone languages, and whether the size of tone inventory has an effect on IF0. Languages with two, three, and four tones are represented. Our results across languages differ, with some languages exhibiting evidence of IF0, while one language shows little or no effect. As it happens, this finding is not tied in a straightforward manner to tone inventory size. We examine whether IF0, where found, is consistently neutralized for Low tone, as suggested in the literature. Again, there is variation across languages, but for the greater part it appears that IF0 is reduced, if not entirely neutralized for L. Third, we explore whether there is evidence for a tone gradient effect on IF0. Again variation exists across languages, but our tentative conclusion in this regard is negative. We are also able to address a number of other issues not specific to tone languages. Our results confirm that there is no difference in IF0 attributable to frontness *vs.* backness of vowels. One of the languages examined has a vowel length contrast in which duration is the distinctive feature, with no corresponding difference in vowel quality. Comparing *F*0 of long and short vowels confirmed that for this language vowel duration does not influence IF0. Finally, for three of the four languages studied, a full range of vowels, from high to low, was included, with the intention of determining whether in these languages there is a vowel gradient with respect to IF0. Our findings do not permit a clear impression in this regard. In presenting our results, two caveats must be borne in mind: first, in hypothesizing that IF0 may be constrained in tone languages, we are asserting the null hypothesis, that there is no difference in *F*0 between high and low vowels. This is unavoidable, but does mean that results failing to demonstrate IF0 must be treated with caution. Second, given the setting of the languages involved, out of necessity we are working with small sample sizes. For these reasons our results are offered as suggestive only.

The four tone languages which contribute new data to the IF0 debate are Ibibio, a Cross River language spoken in SE Nigeria; Kunama, a Nilo-Saharan language spoken in SW Eritrea; Mambila, a Bantoid language spoken in the Nigeria–Cameroon borderland; and Dschang (also known as Bamileke-Dschang, Yemba), a Grassfields Bantu language spoken in the Western Province of Cameroon. They are all "register" tone languages (Pike, 1948), that is, characterized by systems of level tones. Further information on their vowel and tone inventories is given below in the relevant sections.

2. Review of previous work

Little work investigating IF0 has been done using Africa languages. In the survey of Whalen and Levitt, only three are mentioned: Yoruba (Hombert, 1977), Iṣẹkiri

(Ladefoged, 1968); and Hausa (Pilszczikowa-Chodak, 1972). In this section, we review this work and bring to the discussion additional research on Yoruba as well as work on Defaka and Chumburung.

2.1. *Hausa*

Hausa (Chadic, N. Nigeria) has two level tones, though from the data presented in Pilszczikowa-Chodak (1972) only High tone can be examined; only single repetitions from a single male speaker of each item were recorded/measured, although different vowels are repeated. So, while IF0 appears to exist for H in Hausa ([i] = 141 Hz, [u] = 122 Hz, [a] = 117 Hz), it is not possible to ascertain whether IF0 is neutralized with Low tone. It is also worth drawing attention to one further aspect of the Hausa data, that the mid vowel /o/, at 112 Hz, unexpectedly has a lower *F*0 than /a/.

2.2. *Iṣẹkiri*

Iṣẹkiri (Benue-Congo, S. Nigeria) has three level tones. Whalen and Levitt, drawing on information from Ladefoged (1968) report an IF0 difference of 5 Hz; this pertains to both High-toned /i/ and /u/ compared to the average *F*0 of the carrier phrase used. Mid- and Low-toned /i/ and /u/ averaged 4 Hz greater than the carrier phrase. A number of observations may be made, bearing in mind that only one speaker was tested in Ladefoged's study. First, only the high vowels were reported to demonstrate IF0, though all seven Iṣẹkiri vowels were tested. Second, with a difference of only 1 Hz between High tones on the one hand, and Mid and Low tones on the other, IF0, to the extent it can be said to exist, is not neutralized for Low tones, and there is no IF0 gradient for tone in Iṣẹkiri. Note also that the observed effect of 4–5 Hz is small compared to the average IF0 of 15.3 Hz which Whalen and Levitt found across their entire sample.

2.3. *Yoruba*

In addition to Hombert (1977), cited by Whalen and Levitt, three other studies of IF0 in Yoruba (Benue-Congo SW Nigeria), are available. Alo (1990) examined a single male speaker of Tsabe, a variety of Yoruba spoken in Benin Republic, Laniran (1992), a single female speaker and Bakare (1995) presents data from two speakers, one male and one female. Yoruba generally follows the expected trends. Speakers in all four studies show IF0; Hombert and Laniran both include the full range of vowels, with Hombert's results showing the predicted vowel gradient effect, at least loosely (little difference exists between [ɛ, ɔ, a], and with H, [a] is slightly higher than [ɛ, ɔ], though for L [i] actually has the lowest *F*0). For Laniran, the mid vowels do not follow expectations, and her statistics showed only a significant difference between the vowel pairs [i] and [a] and [o] and [a]. Alo's results demonstrate quite clearly the existence of a tonal gradient, showing a difference of 20.9, 8.3 and 1.2 Hz for H, M, and L, respectively. Bakare's study was not aimed at examining IF0 but *F*0 values for /i, u, a/ are included, allowing IF0 to be calculated. Both speakers show IF0 effects for all tones (male: H = 15.3 Hz; M = 10.5 Hz; L = 8.3 Hz; female: H = 31.5 Hz; M = 17 Hz; L = 21 Hz). The Yoruba results are interesting when converted to semitones, permitting averaging across speakers from the four different studies (six speakers, total). These show mean differences of 1.6, 1.1, and 0.8

semitones for H, M, and L, respectively, confirming the existence of a tonal gradient, with an average IF0 across tones of 1.2 semitones.

2.4. *Defaka*

Defaka (Ijoid, SE Nigeria) has two tones, H and L. Information on Defaka is reported by Shryock, Ladefoged & Williamson (1996/97). This report affords only a partial view of intrinsic vowel *F*0, as it presents data only for vowels with low tones, and /u/ is not included since, while it does exist in the language, it did not occur with L in the data collected. (The data were collected primarily for purposes other than examining IF0.) Up to 12 male speakers were recorded, each giving one or two repetitions of the test words, with some of the words containing two tokens of the target vowel. Shyrock *et al.* point out that with one exception (the difference between [o] and [a]), differences between vowel pairs are not significant. Attention should be drawn to the fact that results have been collapsed across speakers probably involving different pitch ranges, and this may have obscured the true picture. Nevertheless, their results do show a tendency for an IF0 effect with L, grouping high vowels /i, e, o/ against the low vowels /ɛ, ɔ, a/.

2.5. *Chumburung*

Chumburung (Kwa, East-Central Ghana) has two tones, H and L, and a system of nine vowels falling into two sets, /i, u, e, o/ and /ɪ, ʊ, ɛ, ɔ, a/, according to [± ATR] harmony. Snider (2001), based on data from three speakers (one female and two male), reports a correlation between vowel height and *F*0, with the difference in *F*0 between high and low vowels statistically significant; however, there is no significant difference in vowel *F*0 across H and L tones, or between front and back vowels. One of the interesting aspects of this study is that it is the first reported in which *F*0 of [+ ATR] vowels may be systematically compared with that of [− ATR] vowels. Although Snider does not address this issue explicitly, values for the two sets of vowels are presented separately. The retracted vowels consistently show a slightly lower *F*0 than their [+ ATR] congeners, though again the differences are not significant (Snider, pers. comm.). A number of possible explanations exist to account for the lower *F*0 of [− ATR] vowels. Ewan (1979) has suggested that constricting the pharynx reduces the tension and increases the thickness of the vocal folds, thereby lowering *F*0 (he discusses the effect with respect to pharyngealized consonants, but the same reasoning applies to [− ATR] vowels). Also, in addition to an expanded pharynx, a [+ ATR] articulation is often accompanied by lowering of the larynx; Ewan (1979) and Rossi and Auteserre (1981) have argued that larynx lowering can have the effect of increasing *F*0. A further possibility is that bunching of the tongue dorsum, resulting in a higher tongue position, is a frequent effect of a [+ ATR] articulation, and this also may instigate a higher *F*0. Finally for Chumburung, when vowel types are combined and *F*0 values converted to semitones and averaged across speakers, the mean difference in *F*0 between high and low vowels in Chumburung is 1.2 semitones.

Examination of IF0 in these five African tone languages suggests an effect which departs to some extent from the general trends noted in Whalen & Levitt (1995). Each of these studies reports an IF0 effect (not necessarily statistically significant, and in some cases no statistics are reported) which, while within the range of variation reported by Whalen and Levitt, is to the lower end of their scale. It is also interesting to observe that in several of the studies which included mid vowels, these vowels often deviated from expectations in either showing an $F0$ higher than the high vowels or lower than the low vowels. It is difficult to assess the importance of these results given the small amount of data provided, but this issue is returned to below.

3. Methodology

There are a number of difficulties associated with experimental work on tone languages (cf. Connell, 2000), not least of which is the problem of finding suitable speakers. The huge majority of tone languages—and certainly those in Africa—are relatively remote geographically, and remain unwritten. As a result, much of the work that has been done on tone languages has concentrated on those few that have substantial urban populations based in academic centers, typically Mandarin, Cantonese, and Thai. Recordings for the present studies, with the exception of a few expatriate residents in the U.K. (the Ibibio speakers and one Kunama speaker), were done in the field. In the case of Mambila, which has just recently been given an orthography, only six people in the village where the work was carried out were sufficiently versed in the orthography to be able to participate reliably in a study of this nature; of these, only four were available at the time of the fieldwork. The Kunama constitute a rural population which was inaccessible at the time of the fieldwork due to a border dispute between Eritrea and Ethiopia; as a result, only those few Kunama speakers with whom contact could be made in the Eritrean capital, Asmara, were available to participate. As much as possible, identical methodologies were used in examining the four languages, to permit the greatest degree of comparability. However, because the data collection was done over a period of 3 years, and in a variety of locations, there are inevitably some differences in methodology across the four languages.

3.1. *Speakers*

Four adult native speakers were used in each study; for Ibibio and Mambila, two male and two female speakers were recorded; for Kunama, results from three males and one female are reported. A second female was recorded for Kunama, but due to a number of systematic mispronunciations, her results were not used. For Dschang all four speakers were male. The small number of speakers used for each language is an obvious limiting aspect of the present study, in particular because the results of Whalen & Levitt (1995) suggest that a sample size of approximately 20 speakers may be needed to bring out the true situation of IF0 in a given language. The small samples used here were unavoidable, for the sorts of reasons discussed in the previous paragraph. However, it also bears mentioning that well over half the studies contributing to the survey by Whalen and Levitt (and therefore our general knowledge and perception of the universality of IF0) used five or fewer speakers. It is in this context that results reported here should be judged.

3.2. *Speech materials*

For each language, the aim was to have naturally occurring words representing all possible vowel $\times$ tone combinations read in a carrier phrase. This was achieved for all languages with the exception of Dschang, for which only one high and one low vowel were used. For both linguistic and logistic reasons, there was some variation in design of speech materials used for the different languages. The details of these differences are described in the relevant section for each language.

3.3. *Recording and processing*

All recordings, whether in the field or in a studio, were done with a digital recorder; in the field this was normally a Sony WM D-7, while for three speakers recorded in London, a Sony DTC ZE700 recorder was used. The Mambila and Dschang speakers were additionally fitted with a laryngograph (electroglottograph), and measurements were made for the Lx signal. The Ibibio, Kunama, and Mambila recordings were processed using Macquirer software (http://www.sciconrd.com) which uses a cepstrum-based $F0$ extraction algorithm. The Dschang recordings were processed using ESPS waves + software.

3.4. *Measurement criteria*

For each language the basic criteria used for taking $F0$ measurements were the same, and followed the guidelines set out by Connell & Ladd (1990). The underlying principle of this methodology is to seek the tonal target, while taking account of well-documented phonetic characteristics of $F0$ realization, such as the perturbations associated with obstruents preceding, and to a lesser extent following, the vowel. Since tone is a primary focus of the study, this strategy is particularly appropriate (more so, for example, than arbitrarily choosing the mid-point or energy peak of the vowel). Since each of the languages studied here has essentially level tones, identifying the tonal target was generally a straightforward matter; i.e., in the huge majority of cases, regardless of language, a substantial portion of the vowel carried a flat $F0$. (This also means that in practice essentially the same results would have been found had the mid-point of the vowel been used.) Exceptions to this were with Low tones which, for Mambila, were often pronounced with a steady fall, making the choice of target more difficult.[2] For Kunama, Low is typically realized with a slight drop in $F0$. In cases where there was no $F0$ plateau, the measurement was taken approximately 50 ms from the end of the vowel in question. This draws attention to a potential drawback to this methodology in the context of IF0 research using tone languages: it is possible for the tonal target to be reached beyond the vowel in question, however, for obvious reasons measurements must be taken during the vowel. The strategy used here may therefore have had the effect of raising actual values for Ls in some cases.

3.5. *Statistical analyses*

Statistical tests consist mainly of contrasts to compare $F0$ values for different vowels and combinations of vowels. In all cases these were separated for tone and, since there were often considerable differences in $F0$ range, for speaker (i.e., not just male *vs.* female). Since the primary claim concerning IF0 involves high *vs.* low vowels, this is our first

[2]It is typical of (African) register tone languages that Low tones are realized with a fall when occurring prepausally or when the word in question is focused. Despite the fact that test items were pronounced in a sentence frame (or indeed perhaps because of it), Low-toned words were often realized in this manner.

comparison; for this, *F*0 values for /i/ and /u/ were combined Previous results have indicated that there is no significant difference between high front and high back vowels; here we extend this finding by testing all front *vs*. all back vowels. Mid-vowels were tested against both high and low vowels in Ibibio and Kunama.

The following sections provide essential background information for each language examined and report the results, language by language.

4. Ibibio

Ibibio is spoken by approximately 2 000 000 people. It has several dialects, which differ in part in their tone systems. Ibibio has two tones, High and Low, with Downstep (here written ꜜ) affecting H tones. With respect to lexical tone, dialect differences consist primarily in that H–L sequences in some dialects correspond to H–HL sequences (i.e., High–Falling) in other dialects, H–ꜜH sequences in some dialects and H–H–L in others. The latter seems to occur only in a relatively small number of words. There is also some variation in vowel systems across dialects (Urua, 2000), however, a core set of six vowels is common to all, viz., /i, u, e, o, a, ɔ/. Words were selected that contained these six vowels, giving all possible combinations of vowels with H and L. The speech materials used were originally designed and recorded for a different purpose; consequently, rather than multiple tokens of the same word for each vowel–tone combination, as used for the languages examined below, different words were used. There was also a different number of tokens of each vowel–tone combination—in no case less than five, typically seven, and in one case nine. Test items were recorded in the carrier phrase ḿbô ____ ǹnɔ̀ (I say ____ for me).

4.1. *Ibibio results*

4.1.1. *High vs. low vowels*

Table 1 presents average *F*0 values for /i, u/ combined, and compared to /a/, giving the difference in both Hz and semitones. High vowels generally show a higher *F*0 than low

TABLE I. Mean *F*0 values for Ibibio high and low vowels with differences between high and low vowels given in Hertz and semitones. Values for high vowels are averaged

Speaker	Tone	i/u	a	IF0 (Hz)	IF0 (ST)
1 (F)	H	302.9	294.7	8.2	0.48
2 (F)		284.4	271.6	12.8	0.79
3 (M)		161.9	150.3	11.6	1.3
4 (M)		139.8	127.3	12.5	1.6
Mean					1.0
1 (F)	L	220.2	221.0	– 0.8	– 0.07
2 (F)		222.9	213.8	9.1	0.72
3 (M)		123.2	116.2	7.0	1.0
4 (M)		114.4	113.2	1.2	0.18
Mean					0.5

TABLE II. Contrast results for comparison of Ibibio mid-vowels /e, o/ *vs.* /i, u/ and *vs.* /a/

Speaker	Tone	i, u, *vs.* e, o	a *vs.* e, o
1	H	F (1, 67) = 0.11, ns	F (1, 67) = 1.418, ns
2		F (1, 66) = 4.236, $p = 0.044$	F (1, 67) = 0.881, ns
3		F (1, 66) = 5.387, $p = 0.023$	F (1, 66) = 0.289, ns
4		F (1, 63) = 5.316, $p = 0.024$	F (1, 66) = 4.489, $p = 0.038$
1	L	F (1, 67) = 1.814, ns	F (1, 66) = 4.418, $p = 0.039$
2		F (1, 66) = 0.674, ns	F (1, 66) = 7.675, $p = 0.007$
3		F (1, 66) = 0.458, ns	F (1, 63) = 6.553, $p = 0.013$
4		F (1, 63) = 1.949, ns	F (1, 63) = 0.498, ns

vowels, suggestive of IF0. Contrasts show this to be significant for three speakers with High tone (S2, F (1, 66) = 4.953, $p = 0.029$; S3, F (1, 66) = 15.727, $p < 0.001$; and S4, F (1, 63) = 19.698, $p < 0.001$; but S1, F (1, 67) = 1.152, ns). For S3, the difference between high and low vowels is also significant for L though reduced compared to that for H (S3, F (1, 66) = 5.548; $p = 0.021$; S1, F (1, 67) = 0.012, ns; S2, F (1, 66) = 2.047, ns; S4, F (1, 63) = 0.177, ns).

4.1.2. *Front vs. back vowels*

Ibibio /u/ tends to have a higher $F0$ than /i/ (see Appendix A), though statistical analysis generally shows the difference to be non-significant. Contrast results show only one significant difference, S3 for H tone: F (1, 66) = 4.052; $p = 0.048$. The general absence of significant differences between front and back vowels with respect to IF0 confirms results reported previously in the literature, as does the aforementioned tendency for /u/ to be somewhat higher than /i/.

4.1.3. *Mid-vowels*

$F0$ values for Ibibio mid-vowels generally fall between those for high and low vowels, though there are several cases where results deviate from expectations (see Appendix A). For instance, for S1, $F0$'s for /e, o, ɔ/ are all higher than that of /i/ with H, and higher than that of /u/ with L, while for S3, /o/ is slightly higher than /i/ with H. On the other hand, /ɔ/ on occasion is lower than /a/ (S2, H tone; S4, L tone) as are /e, o/ (S4, Low tone). Contrasts comparing the mid-vowels /e, o/ with /i, u/ and with /a/ give mixed results (Table II). Only S4 with H tone shows clear evidence of a gradient, with his mid-vowels being significantly different from both high and low.

4.1.4. *Summary*

The Ibibio results conform to most of the predictions associated with IF0. Three of the four Ibibio speakers demonstrate a statistically significant IF0 with H when combined high vowels /i, u/ are contrasted with /a/, while only one speaker showed a similar result for L, indicating that IF0 for this tone is reduced or neutralized. In order to compare across speakers and across tones, differences in $F0$ were converted to semitones (see Table I). Average IF0 in semitones is greater for H than L, at 1.0 and 0.5 semitones, respectively; this relative ranking is also true for three of the four speakers individually, and confirms the expectation of reduced IF0 for L. The Ibibio results likewise agree that although /u/ tends to be higher in *F*0 than /i/, the difference between the two, and front and back vowels generally, is not significant.

5. Kunama

Kunama has a population of approximately 100 000. It has several dialects, though the only two that have been the subject of scholarly study are Barka and Marda. Kunama has three lexical tones, H, M, and L (Connell, Hayward & Ashkaba, 2000) which also combine to form a number of contours, and which demonstrate a certain amount of tone sandhi, similar to that found in Yoruba: Low before High is realized as rising, High before Low is falling. One major restriction on tone distribution is that L is not found in monosyllabic words. These facts were taken into account in establishing a list of speech materials to be used in the experiment. Kunama has five vowels, /i, u, e, o, a/, all of which also occur as long vowels without any perceptible change in vowel quality from short to long. Thirty words were selected, which permitted all possible combinations of tone, vowel quality, and vowel length ($3 \times 5 \times 2$). All words began with a consonant (which reduced the possibility of sandhi effects), and the target vowel–tone combination was always in the first syllable. The target words were read in the carrier phrase ènáꟷnùbè "Did you sayꟷ?" Each speaker read the randomized list of words 10 times, giving 10 tokens of each vowel–tone combination, or 20 tokens each when the vowel duration contrast is disregarded. It was necessary to reject all tokens of short [ō, ā] due to an unexpected tone coarticulation effect internal to the test items used.

5.1. *Kunama results*

5.1.1. *Vowel length*

It has been assumed that when a vowel length contrast is based only on duration, unaccompanied by differences in vowel quality, it should play no role in IF0, since tongue posture will be the same for both long and short congeners. Since to our knowledge this has never been tested directly, we first examined the results to determine whether any influence on I*F*0 could be attributed to vowel length. To test this, the factors vowel and length were combined to create a single variable in an ANOVA run, with the output split by speaker and tone. Scheffé's test of pairwise comparisons was run to determine if this result was due to differences between members of long–short vowel pairs. This was not the case for any speaker. Though all speakers showed a significant effect at each tone level, there were no instances of a significant difference between members of long–short pairs. Significant results from the Scheffé test are presented in Table III. The Kunama results were therefore collapsed across vowel length for further statistical analyses.

5.1.2. *High vs. low vowels*

Means and standard deviations of vowel *F*0's, presented separately by vowel length are given in Appendix A. These values give a clear indication of IF0, as in each case comparing high vowels to low, the former show a higher *F*0. Table IV presents average

TABLE III. Scheffé test results for pairwise comparisons of Kunama vowels, significant results only

Tone	Speaker 1	Speaker 2	Speaker 3	Speaker 4
H		eː – a*; oː – e**; i – o*	aː – uː***; aː – i***; oː – i**; uː – a*; uː – o*; i – a**; i – o**	
M		eː – u*; aː – u*	aː – u**	eː – u***; aː – u***; oː – u**; i – u*; e – u**
L	aː – uː***; aː – i**; u – a*		aː – uː**; uː – o*	iː – aː**; eː – i***; aː – uː*; aː – i***; aː – e***; aː – o**; aː – u*; oː – i***; oː – e*; i – a**

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE IV. Mean $F0$ values for Kunama high and low vowels, with differences presented in Hertz and semitones. Values for high vowels are averaged

Speaker	Tone	i/u	a	IF0 (Hz)	IF0 (ST)
1	H	160.3	153.9	6.4	0.7
2		168.3	156.8	11.5	1.2
3		216.9	194.2	22.7	1.9
4		238.8	231.1	7.7	0.6
Mean					1.1
1	M	136.0	127.9	8.1	1.1
2		132.4	126.5	5.9	0.8
3		162.6	151.2	11.4	1.3
4		192.3	178.3	6.3	1.3
Mean					1.1
1	L	128.0	120.5	7.5	1.0
2		125.6	124.1	1.5	0.2
3		148.2	139.9	8.3	1.0
4		189.4	178.1	11.3	1.1
Mean					0.8

$F0$ values for combined /i, u/ compared to /a/, by speaker and tone, giving differences between both Hz and semitones. All Kunama speakers show a significant difference between high and low vowels with H tone (S1, $F\ (1, 264) = 14.097$, $p < 0.001$; S2, $F\ (1, 261) = 23.748$, $p < 0.001$; S3, $F\ (1, 265) = 81.308$, $p < 0.000$; and S4, $F\ (1, 289) = 9.976$, $p = 0.002$). A significant IF0 also exists for M for all but one speaker (S1, $F\ (1, 264) = 13.517$, $p < 0.001$; S3, $F\ (1, 265) = 12.223$, $p = 0.001$; S4, $F\ (1, 289) = 19.663$, $p < 0.001$; but S2, $F\ (1, 261) = 3.760$, ns), and similarly, the same three speakers also show this for L (S1, $F\ (1, 264) = 19.534$, $p < 0.001$; S3, $F\ (1, 265) = 10.960$, $p = 0.001$; S4, $F\ (1, 289) = 21.310$, $p < 0.001$; but S2, $F\ (1, 261) = 0.429$, ns). This result appears to contradict the claim that L neutralizes IF0; however, the figures in Table IV, both in Hz and in semitones, confirm that for two speakers (S2, S3) IF0 is reduced for L, if not neutralized, and that there is a gradient from H to M to L, though the mean values in semitones show only a reduction in IF0 for L, with no gradient.

5.1.3. *Front vs. back vowels*

For Kunama, the average $F0$ of /i/ and /u/ across speakers and tones is virtually identical (166 *vs.* 167 Hz). Individually, two speakers (S1 with H and S4 with all tones) do show a significant difference (S1, H: $F\ (1,264) = 9.756$, $p = 0.002$; S4, H, $F\ (1,289) = 5.894$, $p = 0.016$; M, $F\ (1,289) = 8.169$, $p = 0.005$; L, $F\ (1,280) = 4.704$, $p = 0.031$). There are similar instances when all front vowels are contrasted with all back vowels, i.e., /i. e, *vs.* u, o/: (S2, H: $F\ (1,261) = 23.547$, $p = 0.001$; S3, H: $F\ (1,265) = 8.540$, $p = 0.004$; S4, M: $F\ (1,289) = 6.914$, $p = 0.009$; $F\ (1,289) = 5.239$, $p = 0.023$). However, clearly there is no consistent tendency across speakers for a difference between front and back vowels.

5.1.4. *Mid-vowels*

As with Ibibio, $F0$ values for Kunama mid-vowels generally fall between those of high and low vowels, though there are exceptions. Unlike Ibibio, all exceptions involve a mid-vowel being higher than /i/ or /u/. Contrasts comparing /e, o/ with /i, u/ and /e, o/ with /a/ (Table V) suggest a vowel gradient in certain cases, though in others mid-vowels group with the high vowels.

5.1.5. *Summary*

Kunama shows stronger evidence of IF0 than Ibibio. All speakers demonstrate IF0 comparing high and low vowels with H tone, and IF0 also appears with M and L tones for most speakers. Overall, the degree of IF0 is only somewhat reduced for L and no strong evidence exists for an IF0 gradient by tone, with M showing the same degree of IF0 as H. No difference exists in Kunama between front and back vowels, and no strong evidence was found for an IF0 gradient according to vowel height.

TABLE V. Contrast results for comparisons of Kunama mid-vowels with high and low vowels

Speaker	Tone	i, u, *vs.* e, o	a *vs.* e, o
1	H	$F\ (1,264) = 0.040$, ns	$F\ (1,264) = 12.902$, $p = 0.001$
2		$F\ (1,261) = 0.003$, ns	$F\ (1,261) = 23.991$, $p = 0.001$
3		$F\ (1,265) = 49.873$, $p = 0.001$	$F\ (1,265) = 10.569$, $p = 0.001$
4		$F\ (1,289) = 0.069$, ns	$F\ (1,289) = 11.380$, $p = 0.001$
1	M	$F\ (1,264) = 7.494$, $p = 0.007$	$F\ (1,264) = 2.754$, ns
2		$F\ (1,261) = 3.069$, ns	$F\ (1,261) = 0.444$, ns
3		$F\ (1,265) = 5.793$, $p = 0.017$	$F\ (1,265) = 2.927$, ns
4		$F\ (1,289) = 13.051$, $p = 0.001$	$F\ (1,289) = 3.141$, ns
1	L	$F\ (1,264) = \ = 0.991$ ns	$F\ (1,264) = 13.009$, $p = 0.001$
2		$F\ (1,261) = 2.325$, ns	$F\ (1,261) = 3.693$, ns
3		$F\ (1,265) = 1.483$, ns	$F\ (1,265) = 5.366$, $p = 0.021$
4		$F\ (1,289) = 6.572$, $p = 0.011$	$F\ (1,289) = 6.366$, $p = 0.012$

6. Mambila

Mambila (Bantoid, Niger-Congo) is spoken in the Nigeria–Cameroon borderland by approximately 90 000 people. It comprises two large dialect clusters, both of which contain substantial variation. The dialect of Mambila studied here (Ba-Mambila; see Connell, 1999*a*, *b*, 2000) is that of the village of Somié in Cameroon. It has four-level lexical tones, usually labeled 1 (High) to 4 (Low) in the literature; here, we adopt the labels H (High), UM (Upper Mid), LM (Lower Mid) and L (Low) to facilitate cross-linguistic comparison. Ba-Mambila has seven vowels, /i, e, a, ə, ɔ, o, u/, occurring in open syllables. Speech materials used consisted of all possible vowel and tone combinations using naturally occurring words of CV shape (i.e., 28 test words). One exception to this was tòr, 'hill'; no low-toned example of [o] in a CV word could be found. The carrier sentence used, "túé__wágácén", "say__quickly", was comprised entirely of High tones. This introduced a certain amount of tone sandhi, particularly with L targets, and to a lesser degree LM, however, this generally was not a problem with respect to measurement; more importantly, it left H test items, where IF0 is most expected to appear, free of any such influences. A number of tokens were identified as mispronunciations and excluded from the analysis.[3] Exclusions were few; notable ones are for Speaker 1, where all tokens of /i/ on UM were eliminated, and for Speaker 3 a number of LM tokens were discarded.

6.1. *Mambila results*

6.1.1. *High vs. low vowels*

Means and standard deviations for Mambila are given in Appendix A, by speaker and tone. Table VI presents $F0$ values for averaged /i, u/ compared to /a/ for Mambila, giving the difference between high and low vowels in both Hz and semitones. For H, it is notable that only two of the four speakers show a difference in $F0$ that goes in the direction of IF0, and for only one of these is the result significant (S4, $F\ (1,99) = 8.608$, $p = 0.004$; S1, $F\ (1,108) = 3.184$, $p = 0.077$, ns; S2, $F\ (1,105) = 0.034$, $p = 0.854$, ns; S3, $F\ (1,95) = 0.045$, $p = 0.833$, ns). The other two speakers actually show a slight negative difference for this tone.

With the Upper–Mid tone, a similar situation exists; again differences between high and low vowels, are generally small, but may be said to go in the direction of IF0. Only one speaker shows a significant result (S3, $F\ (1,95) = 9.012$, $p = 0.003$; S1, $F\ (1,96) = 1.249$, ns; S2, $F\ (1,105) = 2.463$, ns; S4, $F\ (1,99) = 0.006$, ns), and it should be noted that at 16.2 Hz this is a considerably larger difference than shown by this speaker for her other tones, including the H tone. This result may therefore be suspect. Neither the Lower–Mid tone nor the Low tone showed a significant difference between high and low vowels for any speaker (LM: S1, $F\ (1,108) = 0.024$, ns; S2, $F\ (1,105) = 0.307$, ns; S3,

[3] It was not possible to obtain a native speaker's judgments of productions after the recordings were made, and so only those tokens which were obviously mispronunciations, based on the author's 7 years' experience with Mambila, were excluded. These normally involved the substitution of another valid word which differed only in tone and where it was clear, based on the tonal frame of the utterance, that a different word had been inserted. For example, most of our UM exemplars were verbs; in Mambila the imperative of UM verbs is formed by changing the tone to H, and it seemed clear in some instances when the speaker gave the imperative form. Similarly, LM verbs change to L in the imperative.

TABLE VI. Mean $F0$ values for Mambila high and low vowels, with differences presented in Hertz and semitones. Values for high vowels are averaged

Speaker	Tone	i/u	a	IF0 (Hz)	IF0 (ST)
1	H	186.6	181.0	5.6	0.5
2		131.0	131.6	− 0.6	− 0.1
3		233.7	234.8	− 1.1	− 0.1
4		225.2	214.8	10.4	0.8
Mean					0.3
1	UM	155.4	151.4	4.0	0.5
2		119.1	114.0	5.1	0.8
3		210.5	194.3	16.2	1.4
4		200.9	200.6	0.3	0.02
Mean					0.7
1	LM	135.7	135.2	0.5	0.1
2		105.2	103.4	1.8	0.3
3		188.9	186.8	2.1	0.2
4		175.9	173.0	2.9	0.3
Mean					0.2
1	L	98.0	99.8	− 1.8	− 0.3
2		86.9	87.2	− 0.3	− 0.1
3		160.6	165.2	− 4.6	− 0.5
4		158.8	157.4	1.4	0.1
Mean					− 0.2

$F\ (1,95) = 0.125$, ns, S4, $F\ (1,95) = 0.575$, ns. L: S1, $F\ (1,108) = 0.063$, ns; S2, $F\ (1,105) = 0.009$, ns; S3, $F\ (1,95) = 0.268$, ns, S4, $F\ (1,95) = 0.039$, ns).

Mambila, then, shows very little in the way of IF0: only one speaker shows a significant result with H, where IF0 is most expected to appear, while a different speaker does with UM. For both of these, the result stands out from the general tendency to show only a very small difference in $F0$ between high and low vowels.

6.1.2. *Front vs. back vowels*

It is interesting to note that contrary to findings in most other languages, including Ibibio and Kunama, $F0$ means for Mambila /i/ and /u/ (see Appendix A), show /i/ to have a higher $F0$ than /u/ for all speakers, regardless of tone (163 *vs.* 159 Hz). Two sets of contrasts were done to compare $F0$ of front and back vowels. First, /i/ was contrasted with /u/, and second all front vowels were compared to all back vowels (i.e., /i, e, *vs.* u, o, ɔ/). No significant difference was found for any speaker on any tone. One possible explanation for the higher $F0$ of /i/ is that, although Mambila does not have [ATR] vowel harmony, /u/ may in fact be realized with a [− ATR] articulation. This has yet to be confirmed instrumentally, but does fit with auditory impressions.

6.1.3. *Mid-vowels*

As with both Ibibio and Kunama, there are a number of instances where $F0$ of mid-vowels does not always fall between that of high and low vowels; indeed this tendency with Mambila is stronger than with the former two languages. Differences are slight, however, as would be expected given the lack of overall significant difference in the Mambila results. Because there is no difference between high and low vowels, tests were not done for the mid-vowels.

6.1.4. *Summary*

Mambila shows little or no evidence of IF0. *F*0 differences between high and low vowels, with only two exceptions, are not significant and differences in means are considerably lower than those reported both for Ibibio and Kunama, above, and in the literature generally. This is quite clear when the values in Hz are converted to semitones and averaged across speakers (Table VI); for all but one tone (UM), the semitone value is well below half a semitone difference. It is interesting to note though, that despite the general non-significance of the Mambila results, one characteristic at least reminiscent of IF0 does appear, which is that the differences between high and low vowels, such as do exist, are smallest with L, where IF0 is expected to be reduced or neutralized. One plausible interpretation of these results is that "latent" IF0 exists in Mambila, but is constrained to minimize its effect on tone realization. This is discussed further below.

7. Dschang

Dschang has two tones, High and Low, and, unusually, downstep can affect Low as well as High. This leads to eight possible two-tone sequences: HH, H$^{\downarrow}$H, HL, H$^{\downarrow}$L, LH L$^{\downarrow}$H, LL, L$^{\downarrow}$L. Additionally, double downstep (H$^{\downarrow\downarrow}$H) is attested and there are two kinds of phrase-final Low tones (falling and non-falling). Dschang is prominent in the literature for these phenomena (e.g., Tadejeu, 1974; Hyman, 1985; Clark, 1993; Stewart, 1993; Bird, 1994). The dialect of Dschang examined here, spoken in the village of Bafou in Cameroon, has eight vowels /i, ɛ, ɨ, ə, a, u, o, ɔ/ (Bird, 1999); only two of these, [u] and [a], were used in this study.

The speech materials used for the Dschang experiment consisted of VCVC words containing /u/ and /a/ with H, $^{\downarrow}$H, and L tones read in two carrier phrases, lèkàŋ__and lèkàŋ__àzá; "choosing__" and "choosing__POSS", respectively. Adding the possessive marker has the effect of converting $^{\downarrow}$H to $^{\downarrow}$L, thereby giving the four main tones of Dschang. The resulting phrases were read in sequential order, alternating possessive and nonpossessive phrases. A minimum of 10 repetitions was recorded for each speaker. Due to having used only the two vowels, it was not possible to consider differences between front and back vowels, or the place of mid-vowels.

7.1. *Dschang results*

*F*0 means and standard deviations for Dschang are given in Appendix A. These measurements show Dschang to match expectations in that /u/ generally shows a higher *F*0 than /a/. Table VII presents the Dschang results in terms of the differences between these vowels in Hz and in semitones. For H, this difference is significant for three of the four speakers (S1, $F\ (1,39) = 44.134$; $p < 0.001$; S2, $F\ (1,48) = 17.450$; $p < 0.001$; S4, $F\ (1,42) = 16.751$; $p < 0.001$; but S3, ns $F\ (1,39) = 2.365$; $p = 0.132$). For $^{\downarrow}$H, all four speakers show a significant difference (S1, $F\ (1,18) = 34.227$; $p < 0.001$; S2, $F\ (1,18) = 10.079$; $p = 0.005$; S3, $F\ (1,19) = 63.076$; $p < 0.001$; S4, $F\ (1,23) = 7.441$;

TABLE VII. Mean *F*0 values for Dschang /u/ and /a/, tone by tone. Differences presented in Hertz and semitones

Tone	Speaker	u	a	IF0 (Hz)	IF0 (ST)
H	1	141.9	128.2	13.7	1.8
	2	125.9	114.9	11.0	1.6
	3	171.8	167.4	4.4	0.4
	4	133.0	123.2	9.8	1.3
	Mean				1.3
$^{\downarrow}$H	1	127.9	118.9	9.0	1.3
	2	114.4	111.1	3.3	0.5
	3	152.9	145.6	7.3	0.8
	4	121.5	110.5	11.0	1.6
	Mean				1.1
L	1	118.1	102.9	15.2	2.4
	2	101.2	96.9	4.3	0.8
	3	132.8	128.1	4.7	0.6
	4	102.4	102.0	0.4	0.1
	Mean				1.0
$^{\downarrow}$L	1	107.2	100.0	7.2	1.2
	2	104.8	102.2	2.6	0.4
	3	125.1	123.5	1.6	0.2
	4	87.4	87.3	0.1	0.0
	Mean				0.5

$p = 0.012$). Three speakers have a significant difference with L (S1, F $(1, 38) = 114.897$; $p < 0.001$; S2, F $(1, 40) = 31.355$; $p < 0.001$; S3, F $(1, 38) = 93.398$; $p < 0.001$; but S4, ns F $(1, 37) = 0.161$; $p = 0.691$), and for the lowest tone, $^{\downarrow}$L, only one speaker has a significant difference (S1, F $(1, 18) = 25.656$; $p < 0.001$; S2, ns F $(1, 18) = 2.958$; $p = 0.103$; S3, ns F $(1, 19) = 2.978$; $p = 0.101$; S4, ns F $(1, 22) = 0.001$; $p = 0.988$).

When averaged across speakers, mean values in semitones not only show that IF0 is reduced for the L tones, but also appear to bear out the hypothesis of a gradient effect from H to L tones (H > $^{\downarrow}$H > L > $^{\downarrow}$L), though individually none of the four speakers reflect this gradient neatly.

7.1.1. *Summary*

The Dschang data reflect quite clearly the characteristics of IF0. Significant differences (with one exception) in *F*0 exist between high and low vowels for all speakers at the three higher tones. For the lowest tone, $^{\downarrow}$L, only one speaker shows a significant difference. In all cases where difference are not significant, in contrast to many of the Mambila results, the high vowels nevertheless show a higher *F*0 than the low vowels.

8. General discussion

8.1. *IF0 in tone languages*

The results of the investigations presented in the preceding sections generally confirm previous reports in the literature concerning the presence of IF0 in tone languages.

TABLE VIII. Comparison across the four languages. IF0 size in semitones

Tone	Ibibio	Kunama	Dschang	Mambila	Mean
H	1.0	1.1	1.3	0.3	0.9
UM		1.1	1.1	0.7	1.0
LM			1.0	0.2	0.6
L	0.5	0.8	0.5	– 0.2	0.4
Mean	0.8	1.0	1.0	0.3	0.8

Ibibio, Kunama and Dschang all show IF0 for most, if not all, speakers. This is true not only of High tone, where IF0 is expected to be found, but also to a large extent with other, lower, tones. Mambila, however, does not conform to expectations. Not only does Mambila not show significant IF0—only 2/16 *F*0 comparisons between high and low vowels were significantly different, compared to 10/12 for Kunama, 12/16 for Dschang, and 4/8 for Ibibio—but indeed only little more than half of these comparisons are in the direction of IF0 (see Table VI). Moreover, *F*0 differences in Mambila are reduced, not only when compared to the other languages investigated here, but also when compared to the general results presented in Whalen and Levitt (1995). They report an average IF0 of 1.65 semitones across languages; for languages examined in which the sample size was comparable to those discussed here (i.e., 3–5 speakers), the range of IF0 is approximately 0.8–4.7 semitones (see Whalen and Levitt, p. 359, Fig. 1). These values may be compared with semitone values for the four languages investigated here, collated in Table VIII.[4] For these languages, all values fall below the Whalen and Levitt mean, and in fact are well to the lower end of their range. This is also true, though to a lesser extent, of those languages reviewed above for which we were able to calculate *F*0 differences in terms of semitones (viz., Yoruba, Chumburung, both with a mean difference of 1.2 semitones). The low values for Mambila are again striking, as they are below the range found in Whalen and Levitt for studies of a comparable sample size. While these results may reflect the small sample size, since, as Whalen and Levitt point out, about 20 speakers may be necessary to reveal the true picture regarding IF0 in a language, it is nonetheless remarkable that these languages in which an IF0 of a large or even average degree might be detrimental to tonal contrasts should cluster to the lower end of the range of distribution for studies involving languages of a comparable sample size. Indeed, this may be no accident; it may well be the case that in these languages, IF0 is to some extent being constrained. This seems especially true of Mambila which is not surprising, given that other work investigating *F*0 in this language has shown that downtrends such as declination, downdrift, and final lowering are counteracted, reduced, or nonexistent (Connell, 1999).

One way of testing whether IF0 is constrained in tone languages would be to compare the results presented here with those of infant learners of a tone language such as Mambila at the babbling stage, before tone categories have been established. Whalen, Levitt, Hsiao & Smorodinsky (1995) for English and French, and Whalen *et al.* (1995) for Mandarin, have shown that infant learners of these languages at the babbling stage show IF0. If, in the case of a tone language such as those examined here, a larger IF0 was found

[4] The arrangement of tones in Table VIII is not intended to suggest an equivalence between, e.g., Kunama M, Dschang ꜜH, and Mambila UM, but is meant merely to give an ordering from H to L. It is interesting to note, though, that roughly similar differences for H and L tones are found for those three languages showing IF0.

in infant learners than in adult speakers, it would strongly suggest that the development of the linguistic use of tone has the effect of constraining IF0. Such a study would be feasible for Ibibio, Kunama, and Mambila, in that language acquisition appears to proceed in these societies as it does in those where such work has already been conducted—e.g., there is no social constraint against infant vocalization that reduces or inhibits infant babbling (general knowledge of the region suggests the situation is similar for Dschang). Ideally, such a study would be done longitudinally, tracking the same speakers from babbling to full control of tone, or even to postpuberty, given developmental changes. A second means of testing whether IF0 is constrained in tone languages would be to compare results from bilingual speakers of tone and nontone languages. Ba-Mambila speakers, for example, typically also speak French (apart from the older generation), and comparing *F*0 values of their productions in the two languages could provide some insight into this issue. Results from such an experiment would have to be interpreted cautiously as the speakers' degree of fluency in the second language would be difficult to control. Finding a reduced degree of IF0 in the tone language relative to the nontone language would provide some evidence to substantiate the claim that tone languages may constrain IF0, but no difference between the two could simply mean that habits from one language may have been transferred to the second.

One of the key questions posed at the outset of this paper was whether the presence or absence, or degree, of IF0 in a tone language is influenced by the size of the tone inventory. A plausible hypothesis was suggested, that the greater the number of tones—or perhaps more appropriately, tone levels—in a language, the more likely it would be that IF0 would be constrained or perhaps suspended altogether. It is tempting to suggest that the apparent absence of IF0 in Mambila is due to its larger tone inventory, thereby substantiating this hypothesis. This temptation, though, must be measured against the evidence of the other languages examined here. Dschang, which also has four tones, and Kunama with three, showed the greatest degree of IF0, i.e., greater than that found in Ibibio with two tones. Therefore, there would appear to be no direct correlation between the number of tones or tone levels and IF0. Rather than simply the number of tones or tone levels in a language, the nature of its tone system appears to be an important factor. This is suggested by comparing Mambila and Dschang. Mambila not only has four-level lexical tones, but these tones also combine to form at least seven bitonal contours, all of which can occur on monosyllables. (Contour tones in Mambila, as in the vast majority of African languages, are uncontroversially analyzable as sequences of level tones.) Tritonal contours are also attested on monosyllables, both lexically and, in particular, through tone operating at the grammatical level. Some of these contours involve adjacent tones (e.g., UM–H), and therefore are realized through relatively small modulations in *F*0. Dschang also has four tones, albeit of a somewhat different phonological status than those of Mambila, since two of these come about through downstep. Contour tones are found on monosyllables, but these are restricted to a simple rise or fall (i.e., L–H, H–L) which spans a greater distance than many of the contours in Mambila. In addition, contours in Dschang are apparently not as frequently occurring as in Mambila. Tone also operates grammatically in Dschang, but with a different range of functions than in Mambila. Despite their similarities, then, there are important differences between the tone systems of the two languages. Looking further afield, the tone languages included in Whalen & Levitt that showed a high degree of IF0 (i.e., similar in magnitude to that found in nontonal languages) are East Asian languages—mostly varieties of Chinese. Contours of course exist in these languages;

indeed, they are considered to be contour tone languages as opposed to register tone languages. But like Dschang, their contours typically involve considerably greater *F*0 spans than those found in Mambila. In short, it is probably true to say that the tonal space of Mambila is more "crowded", not only than that of Dschang, but also than those of many or indeed most other tone languages. They may well account for the negligible presence of IF0 in the language.

Using languages with more than two tones permitted a test of whether or not an IF0 gradient exists across tone. Given the reports that IF0 is present for vowels on H tone, but is reduced or neutralized for vowels on Low tones, it was suggested that intermediate tones should show an IF0 effect somewhere in between these extremes. Indeed, Whalen & Levitt (1955) suggest that a possible explanation for Kammu L showing IF0 is that it might in fact be realized in the mid region of the speaker's pitch range. The plausibility of this hypothesis was supported in part by the results of Hombert (1977) for Yoruba, and confirmed by other work on Yoruba which appears to show an IF0 gradient for tone, and the suggestion by Ladd & Silverman (1984) that (in nontonal languages) the low end of a speaker's pitch range will tend to neutralize IF0. No clear answer emerges to this question from the two languages examined in the present study (Kunama and Dschang) which contribute relevant information (see Table VIII). Mean results for Kunama show no evidence of a gradient, as M clusters with H. The Dschang results, on the other hand, do show what may be considered to be a gradient, though the step between the lowest and next lowest tones is considerably greater than those between the other tones. What may be common to both Dschang and Kunama, then, is that in both the gap is larger between the lowest tone and the one above it than between the others. Seen in this light, additional support is lent to the speculation of Ladd and Silverman that it is the low end of a speaker's pitch range that will tend to neutralize IF0. In this case, it may be that only the lowest *F*0's use a different mechanism, resulting in a dichotomy rather than a continuum.

For three of the four languages investigated in this study a range of vowels was included, to investigate whether a vowel gradient exists with respect to IF0. Such a finding would provide support for claims that the source of IF0 is in muscular adjustments related to tongue height. No strong evidence is provided either for or against this hypothesis by the present results. While the *F*0 of mid-vowels in each of these languages generally falls between that of high and low vowels, there were also numerous instances of mid-vowel *F*0 falling outside this range, and this is also true in many of the languages reviewed earlier. Different explanations for this phenomenon could be mooted. Lower values for /ɔ/ than /a/ in Ibibio, for example, could be attributed to it having a [− ATR] articulation (cf. Ewan, 1979), however, Ibibio /ɔ/ is not consistently lower than /a/, and an [ATR] difference cannot account for those instances of mid-vowels having higher *F*0 than /i, u/ in any of the languages examined here. The main problem in finding an interpretation, though, is simply that the effects are not consistent across speakers. A larger study, focussing specifically on this question and correlating observed vowel *F*0 with a measure of tongue height, will be needed before a more conclusive answer can be provided.

8.2. *IF0 as a phonetic universal*

Two more general issues arise as a result of the findings presented here, and the interpretation that I*F*0 is restricted at least in Mambila, and possibly even to some extent in the other languages. These are, first, the implications for our understanding of the physiological mechanisms involved in IF0 and, second, what it means to be a phonetic universal.

There is currently a consensus that some form of mechanical coupling between articulatory and phonatory systems, usually a variant of the "tongue pull" hypothesis (see, e.g., Ohala & Eukel, 1987), is responsible for IF0. There is, however, no agreement as to exactly what mechanism is used. Several overviews of laryngeal physiology are available which provide indications as to possible sources of IF0 (Hardcastle, 1976; Sawashima & Hirose, 1983; Honda, 1995; Hirose, 1997; with specific regard to IF0 see Sapir, 1989; Vilkman, Sonninen, Hurme & Körkkö, 1996). It is clear from these that the main factor influencing an increase in *F*0 is contraction of the cricothyroid (CT) and vocalis (VOC) muscles, and that the thyroarytenoid muscle (TA, including VOC), is important in effecting a decrease in *F*0 (Hirano, Ohala & Vennard, 1969; Hirano, Vennard & Ohala, 1970). However, both the CT and TA have also been shown to contribute influences opposite to those just specified (Honda, 1983, 1995). Other muscles are also known or believed to contribute to *F*0 control; the genioglossus (GG) and the sternohyoid (SH) contribute to *F*0 increase and decrease, respectively, and as extrinsic laryngeal muscles it is their contribution that one might consider important with respect to IF0, since they are involved in supralaryngeal articulation. Opposing results have been reported in the literature regarding the direct relevance of CT activity to IF0 (Vilkman, Aaltonen, Raimo, Arajävi and Oksanen (1989); Dyhr (1990) report positive correlations between CT activity and IF0, whereas Whalen *et al.* (1998) contradict this finding). Moreover, a number of recent studies (e.g., Whalen *et al.*, 1998; Herman, Beckman & Honda, 1999) warn against simple interpretations of EMG data with respect to *F*0 control (specifically regarding CT activity, but the recommended caution applies to similar work involving other muscles). The picture emerging from this literature, emphasized by Honda, Hirai, Masaki & Shimada (1999), is one of laryngeal control being a complex mechanism which, while well understood in general outline, requires considerably more research before all the details are known.

If not everything is yet known about laryngeal control, even less is known concerning the physiological control of *F*0 in tone languages. Very few studies have investigated this, and those few (Sagart, Hallé, Boyson-Bardies & Arabia-Guidet (1986); Hallé, Niimi, Imaizumi & Hirose (1990); Hallé (1994) for Mandarin; Erickson (1993) for Thai) have examined East Asian languages. No investigations have been carried out on African tone languages, which are typologically different from the tone languages of Asia. In the latter, pitch movement and direction of movement appear to be of primary importance, as opposed to pitch height in the register languages of Africa (see, e.g., Gandour, 1978, 1983; Connell, 2000 and references therein). While we expect *F*0 control to be similar in these two types of tone language, and also to conform to that described for nontone languages, there is no *a priori* reason to assume that they will be the same in all details. Languages that rely on *F*0 movement (contour tone languages), those that require the achievement and maintenance of specific targets at a range of heights (register tone languages) and those that involve the achievement of targets at specific but relatively infrequent locations in the utterance (nontone and pitch accent languages) likely all involve different degrees of *F*0 control and it is not unreasonable to assume that the mechanisms used to achieve this control should vary accordingly.

Given the lack of consensus on the mechanisms governing IF0, together with our as yet insufficient knowledge of the physiological control of *F*0 in tone languages, especially the register-type languages of Africa, discussion of the implications of the present findings, of how a language such as Mambila constrains a putative universal, might appear to be unwarranted speculation. Nevertheless, if IF0 can be constrained in tone languages, two hypotheses may be presented to account for this. On the one hand, it could be seen as a deliberate suppression, involving active control and based on phonological concerns, e.g., the preservation of tonal contrasts from the effect of vowel intrinsic *F*0. This view resembles the "controlled phonetics" presented in Kingston & Diehl (1994), but it is not the same; it is arguable as to whether constraining a process to preserve a contrast is the same as exploiting one for purposes of contrast enhancement. Moreover, nothing can necessarily be inferred about the status of IF0 in other languages, i.e., there is no implication that it is produced deliberately. On the other hand, constraining IF0 could be seen as a by-product or side-effect of tone production; the physiological mechanism used to achieve the control and precision required for a given tone system may be antagonistic to that which otherwise results in IF0, obscuring or overriding the automaticity of IF0. This, too, has its origin in the need to produce and maintain tonal contrasts, though from a broader perspective than implied by the first alternative. A form of explanation related to this alternative is implicit in the speculation of Ladd & Silverman (1984), discussed earlier, that it is a low *F*0 range that neutralizes IF0, and there is clear evidence that a different mechanism is involved in the control of *F*0 at the low end of a speaker's range, i.e., involving increased SH activity and/or vertical movement of the larynx (Hallé, 1994; Erickson, Honda, Hirai & Beckman, 1995; Honda *et al*., 1999). An explanation which also accords with this second alternative is given by Sapir (1989) to account for the reduced IF0 found in singers, who tend to use less vertical laryngeal movement than do nonsingers in modulating *F*0. Singers do this, not to obscure IF0, but in order to produce and control specific pitch differences. While there is at present insufficient evidence to permit choosing between the alternatives outlined here, they have implications for the second issue mentioned above, what it means to be a phonetic universal.

Much of the debate concerning the universality of IF0 is based on the assumption that for a process or phenomenon to qualify as universal, it must be automatic and occur obligatorily in all languages. We have already indicated earlier that this would appear to be too strict a definition. A phenomenon such as declination is an obvious candidate for universal status. It has been found in all languages in which it has been investigated, seems to be automatic in that it occurs except in clearly specifiable circumstances (i.e., it occurs by "default"), and indeed may not even be species-specific (Hauser & Fowler, 1992). Yet within the terms of the IF0 debate, it is apparently denied this status as it appears to be directly controllable. The first alternative discussed in the previous paragraph is in one respect similar; IF0 will occur naturally, except where phonological concerns are overriding. A view that contends that phonetic universals are strictly obligatory would have to admit that, if IF0 is controlled on this basis, it may not be a universal. The second alternative presents less of a challenge to this strict view of universality: IF0 is assumed to be a consequence of normal vowel articulation, and therefore automatic (and presumably "obligatory"), but it is constrained in some languages because the nature of co-occurring articulations (i.e., it may be useful here to view tone as being articulated rather than "phonated"). The degree of control, and therefore the mechanism of control, required is antagonistic to the occurrence of IF0. IF0 can perhaps be viewed as no less universal for this, as its absence is in a sense accidental.

9. Conclusion

Using languages with different-sized tone inventories, this paper set out to examine a set of questions pertaining to IF0 in tone languages. The central question, whether inventory size is a constraining factor on IF0, presupposed that IF0 does exist in tone languages. Results may be interpreted as showing that IF0 may be constrained in tone languages, even to the point where it can be said not to occur. All of the languages investigated in this study showed a smaller than average degree of IF0 and one, Mambila, apparently showed no IF0. However, the size of tone inventory alone apparently cannot account for the reduction or disappearance of IF0. Rather, the nature of the tone system, and specifically the degree of *F*0 modulation used in producing tonal contrasts (to some extent a consequence of inventory size), appears to be the primary factor. The languages studied here also appear to indicate that, while IF0 is frequently reduced or neutralized with L tones, a tonal gradient with respect to IF0 need not be present in languages with multiple tone levels. This is in accord with physiological studies which indicate different mechanisms to be used for *F*0 control in a high as opposed to a low *F*0 range. Three other questions, not specifically related to IF0 in tone languages, were also addressed. It was confirmed that no difference in IF0 should be expected between front and back vowels, even when this is extended to include more than just high vowels. Duration was also shown not to influence the intrinsic *F*0 of vowels in the absence of corresponding differences in vowel quality. Finally, the status of mid-vowels, and whether a vowel gradient exists for IF0 was also examined, though with equivocal results. A much larger study, directed specifically at this question should be undertaken. Indeed, the small number of speakers on which our results are based and the fact that, in suggesting IF0 may be constrained tone languages, we are in essence asserting the null hypothesis, pointing to the tentative nature of all our conclusions. It is hoped that more extensive work, particularly on Mambila, will be possible in the future.

I am grateful to the many people who assisted in this project. In Cameroon, to the Mambila friends for their participation and hospitality, to David Zeitlyn, and to SIL colleagues. Special thanks to Steven Bird for graciously making available the entirety of his Dschang data and for his collaboration on an early version of this work. Thanks to Emma Connell, John Ekpenyong, Uwem Ite, and Eno Urua for the Ibibio data, and to John Harris for assistance with some of the recordings. Thanks to Macca Teclehaimanot, John Abraha Ashkaba, Andrea, and Ambrogio for their participation. Thanks to Mirka Ondrak and Oren Amitay for advice with statistics, and to Bernard Howard for use of the facilities at the Phonetics Laboratory, School of Oriental and African Studies in London. Thanks to Randy Diehl, Kiyoshi Honda, D. H. Whalen and an anonymous reviewer for helpful comments and discussion. Preliminary and partial versions of this work were presented at a number of venues; comments received on those occasions are appreciated. This project was supported in part by ESRC grant R000235283 to the author. Research in Cameroon was conducted under the auspices of the Ministry of Scientific and Technical Research of the Government of Cameroon.

References

Alo, P. O. (1990) Interaction between segments and tone in Tsabe, *Pholia*, **5,** 7–15.

Bakare, C. (1995) Discrimination and identification of Yoruba tones: perception experiments and acoustic analysis. In *Language in Nigeria: essays in honour of Ayo Bamgbose* (K. Owolabi, editor), pp. 32–67. Ibadan: Group Publishers.

Bird, S. (1994) Automated tone transcription. *Proceedings of the first meeting of the ACL special interest group in computational phonology*, Association for Computational Linguistics, pp. 1–12.

Bird, S. (1999) Dschang syllable structure. In *The syllable: views and facts* (H. van der Hulst & N. A. Ritter, editors), pp. 447–476. Berlin: Mouton de Gruyter.

Bolinger, D. (1964) Intonation: around the edge of language, *Harvard Educational Review*, **34,** 282–296.

Clark, M. (1993) Representation of downstep in Dschang Bamileke. In *The phonology of tone—the representation of tonal register* (H. van der Hulst & K. Snider, editors), pp. 29–73. Berlin: Mouton de Gruyter.
Connell, B. (1999) Four tones and downtrend: a preliminary report on pitch realization in Mambila. In *New dimensions in African linguistics and languages. Trends in African linguistics* (P. F. A. Kotey, editor), Vol. 3, pp. 75–88. Trenton, NJ: Africa World Press.
Connell, B. (2000) The perception of lexical tone in Mambila, *Language and Speech*, **43,** 163–182.
Connell, B. A., Hayward, R. J. & Ashkaba, J. A. (2000) Observations on Kunama tone, *Studies in African Linguistics*, **29**(1), 1–41.
Connell, B. & Ladd, D. R. (1990) Aspects of pitch realization in Yoruba, *Phonology*, **7**(1), 1–29.
Diehl, R. L. (1991) The role of phonetics within the study of language, *Phonetica*, **48,** 120–134.
Diehl, R. L. & Kluender, K. R. (1989) On the objects of speech perception, *Ecological Psychology* **1,** 121–144.
Dyhr, N. (1990) The activity of the cricothyroid muscle in the intrinsic fundamental frequency in Danish vowels, *Phonetica*, **47,** 141–154.
Erickson, D. (1993) Laryngeal muscle activity in connection with Thai tones, *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, **27,** 135–149.
Erickson, D., Honda, K., Hirai, H. & Beckman, M. E. (1995) The production of low tones in English intonation, *Journal of Phonetics*, **23,** 179–188.
Ewan, W. G. (1979) Can vowel intrinsic *F*0 be explained by source/tract coupling?, *Journal of the Acoustical Society of America*, **66,** 358–362.
Fischer-Jorgensen, E. (1990) Intrinsic *F*0 in tense and lax vowels with special reference to German, *Phonetica*, **47,** 99–140.
Fowler, C. A. & Brown, J. M. (1997) Intrinsic *F*0 differences in spoken and sung vowels and their perception by listeners, *Perception & Psychophysics*, **59,** 729–738.
Gandour, J. T. (1974) Consonant types and tone in Siamese, *Journal of Phonetics*, **2,** 337–350.
Gandour, J. T. (1978) The perception of tone. In *Tone: a linguistic survey* (V. Fromkin, editor), pp. 41–76. New York: Academic Press.
Gandour, J. T. (1983) Tone perception in far Eastern languages, *Journal of Phonetics*, **11,** 149–175.
Hallé, P. A. (1994) Evidence for tone-specific activity of the sternohyoid muscle in Modern Standard Chinese, *Language and Speech*, **37,** 103–123.
Hallé, P. A., Niimi, S., Imaizumi, S. & Hirose, H. (1990) Modern Standard Chinese 4 tones: EMG and acoustic patterns revisited, *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, **24,** 41–58.
Hardcastle, W. (1976) *Physiology of speech production*. New York: Academic Press.
Hauser, M. D. & Fowler, C. A. (1992) Fundamental frequency declination is not unique to human speech: evidence from nonhuman primates, *Journal of the Acoustical Society of America*, **91,** 363–369.
Herman, R., Beckman, M. E. & Honda, K. (1999) Linguistic models of *F*0 use, physiological models of *F*0 control, and the issue of "mean response time", *Language and Speech*, **42,** 373–399.
Hirano, M., Ohala, J. J. & Vennard, W. (1969) The function of the laryngeal muscles in regulating fundamental frequency and intensity of phonation, *Journal of Speech and Hearing Research*, **12,** 616–628.
Hirano, M., Vennard, W. & Ohala, J. J. (1970) Regulation of register, pitch, and intensity of voice, *Folia Phoniatrica*, **22,** 1–20.
Hirose, H. (1997) Investigating the physiology of laryngeal structures. In *The handbook of phonetic sciences* (W. J. Hardcastle & J. Laver, editors), pp. 116–136. Oxford: Blackwell.
Hollien, H., Mendes-Schwartz, A. P. & Nielsen, K. (2000) Perceptual confusions of high-pitched sung vowels, *Journal of Voice*, **14,** 287–298.
Hombert, J.-M. (1974) Universals of downdrift: their phonetic basis and significance for a theory of tone, *Studies in African Linguistics* (Suppl: 5), 169–183.
Hombert, J.-M. (1977) Consonant types, vowel height, and tone in Yoruba, *Studies in African Linguistics*, **8**(2), 173–190.
Hombert, J.-M. (1978) Consonant types, vowel quality, and tone. In *Tone: a linguistic survey* (V. Fromkin, editor), pp. 77–111. New York: Academic Press.
Honda, K. (1983) Relationship between pitch control and vowel articulation. In *Vocal fold physiology* (D. M. Bless & J. H. Abbs, editors), pp. 286–297. San Diego: College-Hill Press.
Honda, K. (1995) Laryngeal and extra-laryngeal mechanisms of *F*0 control. In *Producing speech: contemporary issues: for Katherine Safford Harris* (F. Bell-Berti & L. J. Raphael, editors), pp. 215–232. New York: AIP Press.
Honda, K., Hirai, H., Masaki, S. & Shimada, Y. (1999) Role of vertical larynx movement and cervical lordosis in *F*0 control, *Language and Speech*, **42,** 401–411.
Hyman, L. M. (1985) Word domains and downstep in Bamileke–Dschang, *Phonology Yearbook*, **2,** 45–83.
Iivonen, A. K. (1989) Regionally determined realization of the standard German vowel system, *Mimeographed Series of the Department of Phonetics, University of Helsinki*, **15,** 21–28.
Inkelas, S. & Leben, W. R. (1990) Where phonology and phonetics intersect: the case of Hausa intonation. In *Papers in laboratory phonology I: between the grammar and physics of speech* (M. E. Beckman & J. Kingston, editors), pp. 17–34. Cambridge: Cambridge University Press.

Jang, T.-Y. (2000) *Phonetics of segmental F0 and machine recognition of Korean speech*. Unpublished PhD dissertation, University of Edinburgh.
Jun, S.-A. (1993) *The phonetics and phonology of Korean prosody*. Unpublished PhD dissertation, The Ohio State University.
Jun, S.-A. (1996) Influence of microprosody on macroprosody, *UCLA Working Papers in Phonetics*, **92,** 97–116.
Kingston, J. (1993) The phonetics and phonology of perceptually motivated articulatory covariation, *Language and Speech*, **35,** 99–113.
Kingston, J. & Diehl, R. (1994) Phonetic knowledge, *Language*, **70,** 419–454.
Ladd, D. R. (1984) Declination: a review and some issues, *Phonology Yearbook*, **1,** 53–74.
Ladd, D. R. (1996) *Intonational phonology*. Cambridge: Cambridge University Press.
Ladd, D. R. & Silverman, K. E. A. (1984) Vowel intrinsic pitch in connected speech, *Phonetica*, **41,** 31–40.
Ladefoged, P. (1968) *A phonetic study of west African languages: an auditory-instrumental survey*. Cambridge: Cambridge University Press.
Laniran, Y. (1992) *Intonation in a tone language: the phonetic implementation of tone in Yoruba*. Unpublished PhD dissertation, Cornell University.
Lindau, M. (1986) Testing a model of intonation in a tone language, *Journal of the Acoustical Society of America*, **80**, 757–764.
Ohala, J. J. & Eukel, B. W. (1987) Explaining the intrinsic pitch of vowels. In *In honor of Ilse Lehiste* (R. Channon & L. Shockey, editors), pp. 207–215. Dordrecht: Foris.
Ohde, R. N. (1984) Fundamental frequency as an acoustic correlate of stop consonant voicing, *Journal of the Acoustical Society of America*, **75,** 224–230.
Pierrehumbert, J. B. (1980) The phonology and phonetics of English intonation. Unpublished Ph.D. dissertation. Massachusetts Institute of Technology.
Pilszczikowa-Chodak, N. (1972) Tone-vowel height correlation and tone assignment in the patterns of verb and noun plurals in Hausa, *Studies in African Linguistics*, **3,** 399–421.
Pike, K. L. (1948) *Tone languages*. Ann Arbor: University of Michigan Press.
Reinholt Peterson, N. (1978) Intrinsic fundamental frequency of Danish vowels, *Journal of Phonetics*, **6,** 177–189.
Rossi, M. & Auteserre, D. (1981) Movements of the hyoid and the larynx and the intrinsic frequency of vowels, *Journal of Phonetics*, **9,** 233–249.
Sagart, L., Hallé, P., Boysson-Bardies, B. & Arabia-Guidet, C. (1986) Tone production in Modern Standard Chinese: an electromyographic investigation, *Cahiers de Linguistique Asie Orientale*, **15,** 153–174.
Sapir, S. (1989) The intrinsic pitch of vowels: theoretical, physiological, and clinical considerations, *Journal of Voice*, **3,** 44–51.
Sawashima, M. & Hirose, H. (1983) Laryngeal gestures in speech production. In *The production of speech* (P. F. MacNeilage, editor), pp. 11–38. New York: Springer-Verlag.
Shryock, A., Ladefoged, P. & Williamson, K. (1996/97) The phonetic structures of Defaka, *Journal of West African Languages*, **26**(2), 3–27.
Snider, K. (2001) Pitch and vowel quality in Chumburung. Paper presented at the *32nd Annual Conference on African Linguistics*, Berkeley, March 2001.
Stewart, J. M. (1993) Dschang and Ebrie as Akan-type total downstep languages. In *The phonology of tone: the representation of tonal register* (H. v. d. Hulst & K. Snider, editors), pp. 185–244. Berlin: Mouton de Gruyter.
Tadajeu, M. (1974) Floating tones, shifting rules, and downstep in Dschang–Bamileke, *Studies in African Linguistics* (Suppl: 5), 283–290.
Ternström, S., Sundberg, J. & Colldén, A. (1988) Articulatory *F*0 perturbations and auditory feedback, *Journal of Speech and Hearing Research*, **31,** 187–192.
Urua, E.-A. E. (2000) *Ibibio phonetics and phonology*. Cape Town: Centre for Advanced Studies of African Societies.
Vilkman, E., Aaltonen, O., Raimo, I., Arajävi, P. & Oksanen, H. (1989) Articulatory hyoid-laryngeal changes vs cricothyroid muscle activity in the control of intrinsic *F*0 of vowels, *Journal of Phonetics*, **17,** 193–203.
Vilkman, E., Sonninen, A., Hurme, P. & Körkkö, P. (1996) External laryngeal frame function in voice production revisited: a review, *Journal of Voice*, **10**(1), 78–92.
Whalen, D. H., Gick, B., Kumada, M. & Honda, K. (1998) Cricothyroid activity in high and low vowels: exploring the automaticity of intrinsic *F*0, *Journal of Phonetics*, **27,** 125–142.
Whalen, D. H. & Levitt, A. G. (1995) The universality of intrinsic *F*0 of vowels, *Journal of Phonetics*, **23,** 349–366.
Whalen, D. H., Levitt, A. G., Hsiao, P.-L., Irwin, J. & McGowan, W. (1995) Intrinsic *F*0 in the babbling of Mandarin-learning infants. In *Proceedings of the 13th international congress of phonetic sciences* (K. Elenius & P. Branderud, editors), Vol. **3,** pp. 420–423. Stockholm: ICPhS.
Whalen, D. H., Levitt, A. G., Hsiao, P.-L. & Smorodinsky, I. (1995) Intrinsic *F*0 of vowels in the babbling of 6-, 9-, and 12-month-old French and English-learning infants, *Journal of the Acoustical Society of America*, **97,** 2533–2539.

Appendix A: Means and standard deviations of intrinsic vowel *F*0

A.1. *Ibibio* (*Tables AI–AIV*)

TABLE AI. Means and standard deviations for Ibibio Speaker 1 (F)

	i	e	a	ɔ	o	u
H	297.0 (7.9)	301.6 (5.8)	294.7 (6.7)	305.5 (3.6)	305.6 (7.1)	308.8 (7.8)
L	219.3 (4.5)	229.3 (5.4)	221.0 (7.1)	224.0 (8.7)	228.0 (4.9)	221.0 (0.4)

TABLE AII. Means and standard deviations for Ibibio Speaker 2 (F)

	i	e	a	ɔ	o	u
H	278.0 (7.3)	272.1 (4.4)	271.6 (3.9)	269.0 (3.9)	277.0 (3.7)	290.7 (5.9)
L	227.0 (3.9)	228.8 (4.6)	213.8 (5.1)	226.0 (4.9)	224.8 (4.6)	218.8 (4.2)

TABLE AIII. Means and standard deviations for Ibibio Speaker 3 (M)

	i	e	a	ɔ	o	u
H	158.3 (4.5)	152.7 (2.7)	150.3 (1.4)	149.5 (1.9)	159.9 (1.8)	165.5 (5.0)
L	122.0 (0.7)	123.5 (2.1)	116.2 (1.3)	121.1 (2.4)	126.2 (1.4)	124.4 (1.1)

TABLE AIV. Means and standard deviations for Ibibio Speaker 4 (M)

	i	e	a	ɔ	o	u
H	138.2 (2.4)	133.7 (1.5)	127.3 (4.2)	131.5 (2.6)	134.9 (2.4)	141.3 (1.6)
L	113.1 (1.4)	110.7 (1.7)	113.7 (2.4)	110.8 (1.1)	112.0 (1.5)	115.6 (2.6)

A.2. *Kunama* (*Tables AV–AVIII*)

TABLE AV. Means and standard deviations for Kunama Speaker 1 (M)

Tone	Vowel length	i	e	a	o	u
	l	158.6 (6.8)	165.1 (7.3)	154.1 (7.4)	158.00 (7.16)	160.5 (7.1)
H	*s*	155.7 (6.8)	161.7 (11.0)	153.6 (7.3)	154.9 (5.6)	166.0 (7.2)
	Mean	157.2 (6.8)	163.4 (9.3)	153.9 (7.2)	156.5 (6.4)	163.3 (7.6)
	l	134.6 (4.9)	132.0 (6.7)	127.9 (5.8)	130.7 (4.6)	136.6 (5.3)
M	*s*	137.3 (5.5)	133.4 (7.7)			135.2 (3.7)
	Mean	136.0 (5.3)	132.7 (7.1)	127.9 (5.8)	130.7 (4.6)	135.95 (4.5)
	l	125.2 (4.1)	124.*x* (5.4)	118.9 (3.5)	126.9 (6.1)	131.3 (3.8)
L	*s*	129.4 (4.4)	127.1 (4.8)	122.1 (4.0)	127.5 (5.6)	126.0 (5.5)
	Mean	127.3 (4.7)	126.0 (5.1)	120.5 (4.0)	127.2 (5.8)	128.7 (5.4)

TABLE AVI. Means and standard deviations for Kunama Speaker 2 (M)

Tone	Vowel length	i	e	a	o	u
	l	162.6 (10.2)	173.0 (11.2)	163.6 (9.6)	154.1 (9.8)	166.4 (9.7)
H	*s*	174.7 (13.3)	181.8 (13.5)	150.0 (7.5)	165.2 (13.7)	169.4 (17.5)
	Mean	168.7 (13.1)	177.4 (12.9)	156.8 (11.0)	159.4 (12.8)	167.9 (13.9)
	l	133.6 (4.1)	126.6 (4.7)	126.5 (4.0)	130.0 (2.9)	129.1 (5.6)
M	*s*	131.2 (3.6)	127.9 (4.4)			136.1 (7.5)
	Mean	132.4 (4.0)	127.3 (4.5)	126.5 (4.0)	130.0 (2.9)	132.4 (7.3)
	l	125.3 (5.1)	129.8 (3.9)	123.7 (3.2)	128.5 (4.6)	128.8 (5.7)
L	*s*	125.7 (6.4)	126.1 (2.6)	124.4 (3.4)	129.9 (5.8)	122.3 (4.5)
	Mean	125.5 (5.6)	128.0 (3.8)	124.1 (3.2)	129.2 (5.1)	125.7 (6.0)

TABLE AVII. Means and standard deviations for Kunama Speaker 3 (M)

Tone	Vowel length	i	e	a	o	u
	l	211.0 (6.1)	206.5 (5.5)	191.5 (6.6)	198.7 (6.6)	221.2 (9.6)
H	*s*	225.0 (27.7)	207.8 (7.9)	196.8 (14.2)	196.3 (8.9)	210.1 (8.8)
	Mean	218.0 (20.8)	207.2 (6.6)	194.1 (11.1)	197.5 (7.7)	215.6 (10.6)
	l	163.0 (5.3)	156.8 (5.8)	151.2 (6.4)	158.4 (9.1)	160.4 (6.5)
M	*s*	161.9 (8.6)	154.5 (6.6)			164.9 (8.0)
	Mean	162.5 (6.9)	155.7 (6.2)	151.2 (6.4)	158.4 (9.1)	162.6 (7.5)
	l	148.5 (4.7)	146.9 (4.2)	138.8 (2.6)	145.5 (4.1)	152.2 (7.4)
L	*s*	149.1 (11.5)	141.4 (5.6)	140.9 (4.4)	148.9 (4.6)	142.9 (5.6)
	Mean	148.8 (8.5)	144.2 (5.6)	139.9 (3.7)	147.2 (4.7)	147.6 (8.0)

TABLE AVIII. Means and standard deviations for Kunama Speaker 4 (F)

Tone	Vowel length	i	e	a	o	u
	l	234.9 (11.7)	234.8 (11.7)	232.0 (9.4)	236.0 (10.2)	239.9 (9.0)
H	*s*	235.8 (7.8)	247.6 (9.7)	230.2 (11.1)	238.8 (7.6)	244.5 (11.7)
	Mean	235.4 (9.7)	241.2 (12.3)	231.1 (10.1)	237.4 (8.9)	242.2 (10.5)
	l	189.56 (13.4)	179.8 (7.6)	178.3 (8.1)	185.9 (8.1)	190.5 (10.4)
M	*s*	186.7 (7.8)	185.1 (5.1)			204.8 (7.5)
	Mean	188.1 (10.8)	182.3 (7.0)	178.3 (8.1)	185.9 (8.1)	196.5 (11.6)
	l	187.3 (8.0)	180.6 (8.1)	173.6 (5.7)	177.7 (5.2)	186.6 (7.0)
L	*s*	197.6 (6.4)	191.0 (6.4)	182.7 (5.6)	187.8 (7.6)	180.0 (5.9)
	Mean	192.4 (8.8)	185.8 (8.9)	178.1 (7.2)	182.8 (8.2)	186.3 (6.3)

A.3. *Mambila* (*Tables AIX–AXII*)

TABLE AIX. Means and standard deviations for Mambila Speaker 1 (M)

Tone	i	e	ə	a	ɔ	o	u
H	186.8 (8.4)	193.0 (6.4)	184.6 (8.1)	181.0 (5.0)	182.0 (4.1)	189.8 (7.1)	186.3 (7.2)
UM	—	155.4 (4.9)	157.2 (4.3)	151.4 (6.7)	156.2 (5.1)	160.2 (1.9)	155.4 (6.6)
LM	137.4 (6.0)	134.2 (2.9)	143.0 (9.7)	135.2 (5.0)	134.2 (3.9)	134.4 (5.5)	134.0 (6.5)
L	100.6 (6.6)	98.6 (6.9)	103.0 (4.6)	99.8 (4.5)	109.0 (5.6)	104.2 (5.1)	97.4 (3.5)

TABLE AX. Means and standard deviations for Mambila Speaker 2 (M)

Tone	i	e	ə	a	ɔ	o	u
H	133.0 (3.4)	134.6 (3.0)	132.0 (1.6)	131.6 (4.8)	130.4 (2.6)	129.0 (3.1)	129.0 (0.7)
UM	120.8 (9.2)	115.7 (2.3)	113.0 (20.8)	114.0 (3.2)	116.6 (2.5)	112.2 (2.6)	117.4 (1.7)
LM	108.4 (8.5)	110.0 (23.8)	104.4 (4.8)	103.4 (2.1)	113.6 (3.4)	100.2 (6.5)	102.0 (3.5)
L	87.4 (3.0)	86.2 (2.7)	96.5 (8.2)	87.2 (1.5)	89.0 (1.6)	92.4 (1.1)	86.4 (3.1)

TABLE AXI. Means and standard deviations for Mambila Speaker 3 (F)

Tone	i	e	ə	a	ɔ	o	u
H	236.6 (11.9)	241.8 (9.7)	234.2 (3.0)	234.8 (17.9)	233.0 (12.1)	243.0 (9.3)	230.8 (11.1)
UM	213.0 (13.3)	205.2 (13.1)	212.7 (14.0)	194.3 (7.5)	217.0 (14.0)	205.6 (6.5)	208.0 (11.9)
LM	188.2 (5.1)	176.0 (2.8)	163.5 (5.0)	186.8 (8.4)	181.2 (6.6)	165.0 (0.0)	189.5 (3.5)
L	167.6 (7.4)	163.0 (6.0)	156.6 (8.5)	165.2 (6.7)	157.2 (5.7)	173.0 (7.9)	157.4 (7.6)

TABLE AXII. Means and standard deviations for Mambila Speaker 4 (F)

Tone	i	e	ə	a	ɔ	o	u
H	224.8 (6.4)	224.0 (5.9)	227.2 (5.3)	214.8 (6.9)	224.4 (7.5)	227.0 (6.1)	225.6 (7.2)
UM	203.8 (3.4)	199.4 (4.5)	198.2 (5.5)	200.6 (2.0)	204.0 (0.0)	197.0 (4.9)	198.0 (7.5)
LM	176.0 (7.4)	175.5 (3.5)	168.5 (12.9)	173.0 (5.4)	173.2 (3.0)	171.8 (11.7)	175.7 (7.8)
L	160.0 (5.5)	155.2 (3.7)	150.4 (2.0)	157.4 (5.4)	155.2 (3.0)	164.8 (11.4)	157.5 (2.7)

A.4. *Dschang* (*Tables AXIII–AXVI*)

TABLE AXIII. Means and standard deviations for Dschang Speaker 1 (M)

	H	DH	L	DL
/u/	141.9 (5.5)	127.9 (3.6)	118.1 (4.2)	107.2 (2.2)
/a/	128.2 (7.5)	118.9 (3.3)	102.9 (4.8)	100.0 (3.9)

TABLE AXIV. Means and standard deviations for Dschang Speaker 2 (M)

	H	DH	L	DL
/u/	125.9 (10.9)	114.4 (2.0)	101.2 (2.5)	104.8 (3.4)
/a/	114.9 (7.8)	111.1 (2.6)	96.9 (2.44)	102.2 (3.5)

TABLE AXV. Means and standard deviations for Dschang Speaker 3 (M)

	H	DH	L	DL
/u/	171.8 (9.5)	152.9 (1.9)	132.8 (1.6)	125.1 (1.3)
/a/	167.4 (8.6)	145.6 (2.3)	128.1 (1.6)	123.5 (2.8)

TABLE AXVI. Means and standard deviations for Dschang Speaker 4 (M)

	H	DH	L	DL
/u/	133.0 (7.9)	121.5 (13.0)	102.4 (3.5)	87.4 (13.2)
/a/	123.2 (7.9)	110.5 (3.2)	102.0 (1.9)	87.3 (9.2)