
RESEARCH ARTICLE

Intrinsic Fundamental Frequency (IF0) and Tonal Gradient in Yorùbá

Saliu Ayowumi Shittu

Department of English Language and Literature, University of Bahrain, Bahrain

Corresponding Author: Saliu Ayowumi Shittu, **E-mail:** sshittu@uob.edu.bh

ABSTRACT

The present study investigates whether tones in tone languages constrain or control Intrinsic Fundamental Frequency (IF0), similar to the investigation conducted by Hombert (1977) and Connell (2002). Two native speakers of Yorùbá were recorded producing the seven vowels in the language with all three tones. Results of the analysis show that IF0 is found to be a gradient with respect to tones; the High tone possesses a higher F0 than the Mid tone, while the Low tone has the lowest. However, the results provide no strong evidence to support tonal gradient with respect to vowel heights. To some extent, the high vowel [u] appears to follow the prediction regarding tonal gradient, especially in comparison with the low vowel [a]; the other high vowel [i] does not. In most instances, the mid vowels record higher F0 than [i]. These mixed results are possibly due to some extralinguistic factors, such as the speakers' habitual way of speaking, vocal tract shapes, and the larynx, among others.

KEYWORDS

Intrinsic fundamental frequency (IF0), tonal gradient, vowel heights, tone language, Yorùbá.

ARTICLE INFORMATION

ACCEPTED: 25 October 2022

PUBLISHED: 01 November 2022

DOI: 10.32996/ijllt.2022.5.11.5

1. Introduction

Intrinsic fundamental frequency (IF0), also often referred to as intrinsic pitch, is based on the observed tendency for high vowels to have higher fundamental frequency (F0) than low vowels. IF0 is found to be universal across world languages (Whalen, 1995). Based on the assumption that IF0 is a function of vowel height, the high vowels [i] and [u], for instance, are produced with the tongue high in the mouth. These two vowels, also known to have a relatively low first formant (F1), tend to record higher F0 than the low vowels such as [a] and [ɑ]. A study of Danish vowels by Reinholt Petersen (1978) reported that the differences in frequency between the Danish high and low vowels [u] and [a] are between 10 and 30Hz. Other related studies on F0 (e.g., Ogden 2009) observed that female speakers tend to record higher average F0 than male speakers due to the anatomical difference in the construction of the larynx. The typical IF0 range in normal speech for a female speaker ranges between 120-300Hz (Ogden, 2009) or 150-300Hz (Boersma and Weenink, 2015). In comparison, the average F0 for a male speaker ranges between 70-250Hz (Ogden, 2009).

1.1 IF0 in Yorùbá

Yorùbá has three basic tones - High, Mid, and Low. A study of the phonetic interpretation of the three tones by Hayward, Watkins, and Oyetade (2003) found that the High tone is marked by a prominent higher frequency than the other two tones, while the Low tone is marked by a more prominent low frequency. The Mid tone is the least marked of the three. In a related study, Bakare (1995:52), citing F0 as the best differentiator of the Yorùbá tones, revealed that out of the three Yorùbá tones, the High tone has the highest F0. Yorùbá has seven oral vowels. Based on the tongue position during the articulation of the vowels, they are classified as follows (see Table 1). [i] is a high-front vowel, while [u] is also a high vowel but produced further at the back. [e] is a high-mid front vowel, while [o] is a high-mid back vowel. [ɛ], represented as [ɛ̃] in Yorùbá orthography, is a low-mid front vowel, while [ɔ], represented as [ɔ̃] in Yorùbá orthography, is a low-mid back vowel. [a] is a low vowel occupying the central position.

Table 1: Distribution of Yorùbá oral vowels according to tongue position

Tongue position	Yorùbá oral vowels	
	Front	back
high	i	u
high-mid	e	o
low-mid	ɛ	ɔ
low	a	

Studies focusing on IF0 in Yorùbá (e.g., Alo, 1990; Connell, 2002; Hombert, 1977) have revealed disparate results concerning the IF0 gradient with respect to tone. The tonal gradient between the three Yorùbá tones was reported by Alo (1990), who investigated the interaction between vowel quality and tone. The data for the study was sourced from a single male speaker, and it consisted of only the high and the low vowels - [a], [i], and [u]. The results revealed that the F0 difference between high and low vowels is greater with High tones than with Mid and Low tones. Hombert (1977) also investigated IF0 in Yorùbá. Her study suggested that the F0 difference between high and low vowels is more pronounced with High tone than with Low tone. However, the predicted vowel gradient effect was not as conclusive as in Alo (1990). The results did not reveal a significant tonal gradient between the vowels [ɛ], [ɔ], and [a]. Moreover, for the High tone, [a] recorded a slightly higher F0 than [ɛ] and [ɔ], while for the Low tone, [i], a high vowel, recorded the lowest F0. Another related study was Laniran (1992), whose data was obtained from a single female speaker. Her results, especially with respect to the Yorùbá mid vowels, did not follow the predicted tonal gradient effect. The only significant difference observed between vowel pairs was in [i] vs. [a] and [o] vs. [a]. Findings from IF0 studies on other African languages are not inconsistent with those reported above. Connell (2002) investigated the existence of IF0 in four African tone languages - Ibibio, Dschang, Kunama, and Mambila. He found that IF0 is 'frequently reduced or neutralized with Low tones'. However, the study did not conclusively suggest the existence of IF0 in these languages. The results indicated that, to some degree, IF0 is present in three of the four tone languages. The fourth language - Mambila, apparently showed no F0, probably due to the size of the tone inventory or the nature of the tone system. He observed that while IF0 was generally reduced for the Low tone, not all four languages showed the postulated gradient. He further suggested that the existence of a tonal gradient with respect to IF0 may not be necessary for languages with many tone levels.

All these studies suggest that there may be a relationship between IF0 and tone heights. However, while Alo (1990) appears to show clearly that there is a tonal gradient in Yorùbá, the other studies (e.g., Connell, 2002; Laniran, 1992, and Hombert, 1977) do not provide conclusive evidence. The present study seeks to contribute to this scholarship by examining the existence of a tonal gradient in the three Yorùbá tones. Additionally, it will investigate whether there is a tonal gradient with respect to all seven oral vowels in Yorùbá. Unlike most of the previous studies, whose data were obtained from single speakers, the present study will involve two speakers, one male, and one female. Data from the two participants will be compared in light of the areas of the investigation.

The primary expectation for the present study is that the height of the three Yorùbá tones will influence the IF0 range. The high tone is expected to record a higher F0 than the Mid tone. The Low tone is expected to record the lowest IF0 for all the vowels. Regarding tonal gradient with respect to Yorùbá vowels, the study predicts that higher vowels will record higher F0s than mid and low vowels for all three Yorùbá tones.

2. Method

2.1 Participants

To carry out the study, two native speakers of Yorùbá, one male and one female, took part in the experiment. They are both adults in their 40s, and they speak Yorùbá as their first language. They have lived outside the language community for more than 10 years. They speak the Oyo-Ibadan Yorùbá dialect - the mainstream dialect from which standard Yorùbá evolves. Their participation in the study was voluntary.

2.2 Corpus

The corpus consisted of the seven Yorùbá vowels placed post vocally after a voiced bilabial stop [b]. Each vowel was combined with the stop in a word-list reflecting CV tokens. The bilabial stop was used due to two main reasons. First, it has a relatively less important perturbatory effect on the F0 of the following vowel than the velar stops such as [k] and [g] (Hombert and Ladefoged, 1976). Secondly, it provides a wider range of meaningful utterances with the seven Yorùbá vowels when alternated with the three tones, namely High [H], Mid [M], and Low [L]. In situations where the vowel combination with the bilabial stop did not yield meaningful words, nonce words with the appropriate tones were used instead. The original word list (without the tones) is as follows: [ba], [be], [bɛ], [bí], [bo], [bɔ], and [bù]. The final word list, which included the seven Yorùbá vowels produced with the three Yorùbá tones - High (acute accent, H), Mid (unmarked, M), and Low (grave accent, L) -, is presented in Table 2.

Table 2: Vowel combinations with bilabial stops in prevocalic position with their meanings.
 V= vowels; T= tones; H= High tone; M=Mid tone; L=Low tone; NW = nonce words.

V / T	[a]	[e]	[ɛ]	[i]	[o]	[ɔ]	[u]
H	<i>bá</i> 'meet'	<i>bé</i> NW	<i>bẹ</i> 'jump'	<i>bí</i> 'beget'	<i>bó</i> 'peel'	<i>bọ</i> 'fall'	<i>bú</i> 'abuse'
M	<i>ba</i> 'weave'	<i>be</i> NW	<i>bẹ</i> 'peel'	<i>bí</i> 'ask'	<i>bo</i> NW	<i>bọ</i> 'worship'	<i>bu</i> 'rot'
L	<i>bà</i> 'land'	<i>bè</i> NW	<i>bẹ</i> 'beg'	<i>bì</i> 'vomit'	<i>bò</i> 'cover'	<i>bọ</i> 'wear'	<i>bù</i> 'scoop'

2.3 Procedure

The two participants read out the prepared list of stimuli containing the Yorùbá words or nonce words bearing the three tones (Table 1). In total, the word list consisted of 21 tokens (7 vowels x 3 tones). Each token was put in the Yorùbá sentence frame 'Sɔ - ---- soke', meaning: 'Say ---- aloud'. Before the recordings, the participants were informed of instances where the combinations resulted in non-words. In such cases, they were given enough time to practice saying these non-words before the actual recording. Whenever they made any mistakes in pronunciation, the recording was stopped, and that segment was rerecorded. Each token was recorded three times. The recording was made using an AT2020 USB condenser microphone connected directly to the computer for direct recording into the PRAAT program (Boersma and Weenink, 2015). The recordings took place at a quiet place in the participants' homes early in the morning. Each participant produced 63 tokens (7 vowels x 3 tones x 3 repetitions). Thus, for the two participants, a total of 126 tokens were available for analysis.

2.3.1 Acoustic Measurements

The PRAAT shareware (Boersma and Weenink, 2015) was used for the analysis. It gives acoustic displays of the F0 with time-aligned labelling fields. Each segment of the recordings was first annotated into a text grid to identify the vowels for subsequent measurement. The vowel durations were visually identified from the spectrogram as the point following the end of the stop closure and preceding the influence of the initial fricative in 'soke'. Auditory identification was also used to aid visual identification. Figure 1 shows a sample of the spectrogram showing the sentence frame with the text grid for one of the tokens. Following the annotation and identification of the vowel duration, the fundamental frequency (F0) of the vowels for each of the three tones was measured at the center of the vowels.

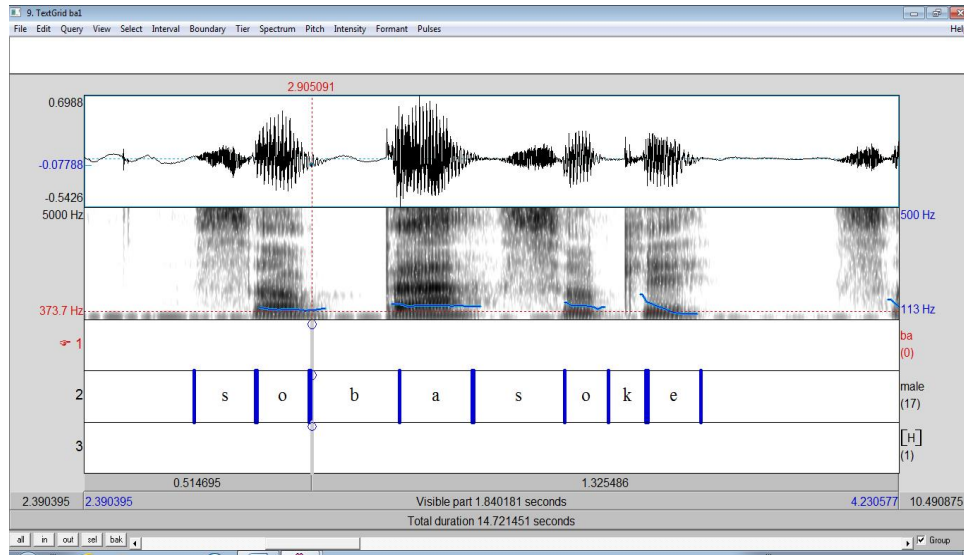


Figure 1: A sample spectrogram showing the sentence frame with the text grid for the male subject.

The three measurements for each token were recorded in a table, and the averages were taken. The averages, presented in Tables 3 and 4, were based on a subset of the three tokens that were recorded for each vowel-tone combination. An overview of the data reveals a relationship between vowel heights and the three tones in Yorùbá. Some of these relationships will be discussed in the results section below.

Table 3: Average F0 readings (in Hz) for the two participants:

T=Tones; V=vowels; M=male, F=female; H=High tone; M=Mid tone; L=Low tone

	T / V	[a]	[ɛ]	[ɔ]	[e]	[o]	[i]	[u]
Male	H	129.2	140.4	140.8	136.6	136.7	129.4	146.6
	M	109	115.2	121.1	116	123.8	113.5	130.1
	L	89.92	102.9	97.73	96.6	103.8	94.3	103
Female	H	211.7	206.7	211.8	219.6	221	215	221.2
	M	184.5	188.1	192	194.6	200.3	190.8	192
	L	150.6	164.7	170.6	172.1	170.6	154.4	167.8

2.3.2 MEL Scale

To get an alternative view of the data, the frequencies were converted from Hertz (Hz) into the MEL scale, which expresses frequency in terms of psychophysical metrics (i.e., perceived pitch). This refers to the perceptual scale of pitches judged by listeners to be equal in distance from one another. The formula used is sourced from tabulated data from Beranek (1949) and is stated below.

$$\text{Mel} = 1127.01048 * \log(f/700 + 1).$$

However, for the purpose of this study, the online version for automatic conversion was employed. As shown in Table 4, the Mel scale allows for better comparisons between the different data.

Table 4: Average Mel Scale reading for the two speakers:
 T=Tones; V=Vowels; M=male; F=female; H=High tone; M=Mid tone; L=Low tone.

Gender	T / V	[a]	[ɛ]	[ɔ]	[e]	[o]	[i]	[u]
Male	H	191	206	207	201	201	191	214
	M	163	172	180	173	184	169	192
	L	136	155	147	146	156	142	155
Female	H	298	292	298	308	309	302	309
	M	264	268	273	276	284	272	273
	L	220	238	246	248	246	225	242

3. Results

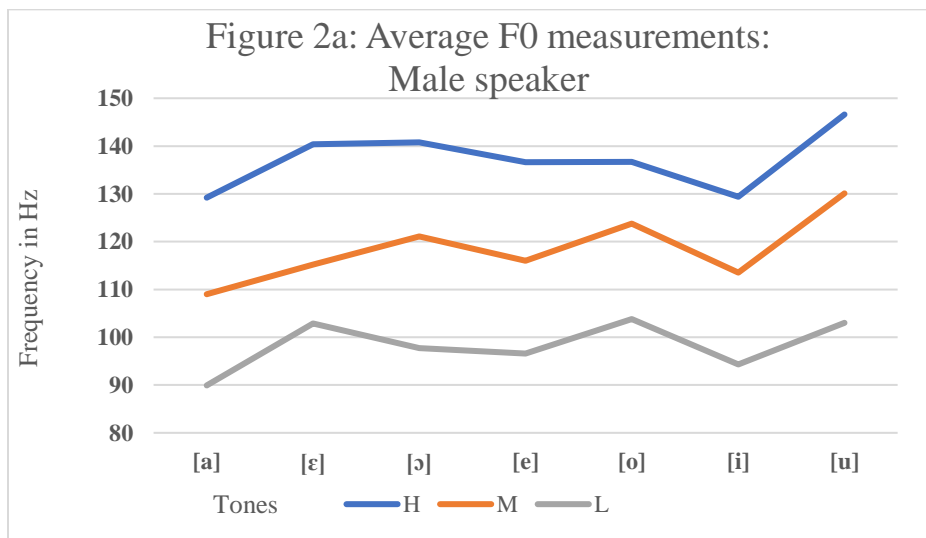


Figure 2a: Average F0 measurements for the Yorùbá vowels with the three tones for the male speaker.

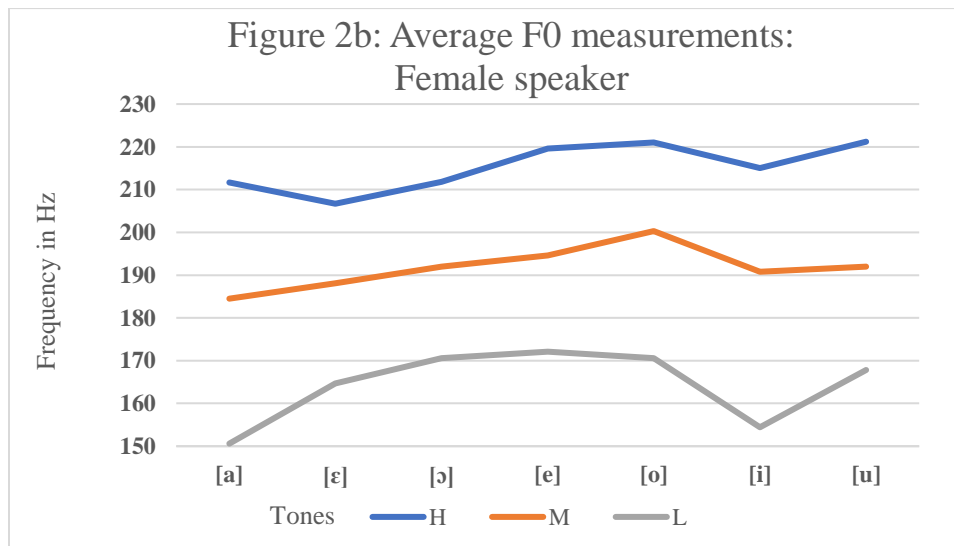


Figure 2b: Average F0 measurements for the Yorùbá vowels with the three tones for the female speaker.

The present study examines the relationship between the three Yorùbá tones and intrinsic fundamental frequency (IF0). The study predicts that there will be a gradient between the three Yorùbá tones. The study also predicts, based on the theory of the universality of IF0, that there will be a tonal gradient based on the height of the seven oral vowels in Yorùbá. Tables 3 and 4, respectively, show the average F0 (Hz) / Mel scale of the seven Yorùbá vowels combined with the three tones, - High, Mid, and Low -, for both the male and the female speakers. These averages are based on a subset of three tokens for each vowel and tone. The results confirm the findings of the previous studies (Alo, 1990; Connell, 2002; Hombert, 1977) with some variations. A general overview of the data in the two tables reveals the following: First, the results confirm the existence of IF0 in Yorùbá, a tone language, in line with Connell (2002), Alo (1990), and Hombert (1977). Second, the IF0 gradient exists with respect to the tones in Yorùbá. The data show that, with the three Yorùbá tones, the F0 magnitude decreases as the tone level moves from High to Low. The F0 readings for all the vowels support this, as indicated in Figures 2a and 2b for both the male and the female subjects, respectively. This aspect of the results supports previous findings by Hombert (1977) and Alo (1990).

IF0 gradient with respect to vowel heights predicts that in a three-level vowel system, mid vowels tend to have an IF0 between that of high and low vowels. With respect to tones, the expectation is that the vowels will show a gradient according to the tone heights. The data from this study present conflicting results. Figures 2a and b indicate the tendency for [u], a high vowel, to have a higher frequency than most of the other vowels for all three tones, although this is more valid for the male speaker than for the female speaker. There is also the tendency for [i]; also a high vowel, to have a generally low frequency than the mid vowels for all three tones, especially for the male speaker. The prediction concerning the tonal gradient effect appears to manifest only in the pair [u] and [a]. For both the male and the female speaker, [u], a high vowel recorded a higher average F0 than [a], a low vowel for the three tones. In contrast, the other high vowel, [i], recorded an average F0 either similar to that of [a] (e.g., for High tone, Male speaker) or just slightly different from that of [a], as shown by the Mid and Low tones for both the male and female speakers. Interestingly, the data revealed that [i] and [a] record the lowest F0 for the three tones for the male speaker, while for the female speaker, the two tones record the lowest F0 only for the Low tone.

The data for the male speaker further reveals that the average F0 for [u] is higher than those of the other vowels for the High and Mid tones. For the Low tone, the average F0 values for [u], [o], and [ɛ] are the same. Moreover, while there are no significant differences in the average F0 values for all the mid vowels [ɛ, ɔ, e, and o] for the three tones, the average F0 value for each of these vowels is higher than that of [i], a high vowel.

By comparison, the readings for the female speaker are even more conflicting regarding tonal gradient. For the three tones, the differences in average F0 values between [i] and [a] are not more than 6Hz. For the High tone, [u], [o], and [e] record average F0 values (219.6 – 221.2Hz), which are higher than all the other vowels. For the Mid tone, [o] has the highest average F0 value. For the Low tone, there are no significant differences in the average F0 values for [o], [ɔ], and [e] in addition to the high vowel [u]. These four vowels record higher F0 values than all the other vowels. Moreover, for the High tone, the average F0s for the high-mid vowels [o] and [e] (221 and 219.6Hz, respectively) are similar to that of the high vowel [u] (221Hz) but higher than the average frequencies recorded for the low-mid vowels [ɛ] and [ɔ] (206.7 and 211.8Hz respectively). However, Mid tone [o] posts an average F0 higher than all the other vowels, including the high vowel [u]. For the Low tone, there are no significant differences in the average values for all the mid vowels (164.7 – 170.6Hz). Their average F0s are similar to that of [u] (167.8Hz) but are higher than that of [i] (154.4Hz) and [a] (150.6Hz).

In sum, the data also support the generalization that female speakers have higher average F0 than male speakers due to the anatomical difference in the construction of the male and female larynx. To some extent, [u] tends to follow the expectation regarding the IF0 gradient, while [i] does not. The mid vowels, on the other hand, show some consistency in recording a higher average F0 than the low vowel. In other instances (mostly for the female speaker), they tend to record F0 values similar to, or higher than, those of the high vowels. The next section will discuss some of the issues raised by the results of the data analysis.

4. Discussion

The data from the two speakers reveal that the three Yorùbá tones follow the expected gradient. The High tone inherently possesses a higher F0 than the Mid tone; the Low tone possesses the lowest F0. What is not proven by the results is the presence of an IF0 gradient with respect to vowel heights. Out of the two high vowels, the data indicate the tendency for [u] to have a higher F0 than [i]. This is possibly due to the active lowering of the larynx while producing the [u] (Reinholt, 1978). Furthermore, when the frequencies for [u] and [i] are contrasted with that of the low vowel [a] for the two speakers, the data indicate that only [u] shows clear evidence of tonal gradient while [i] does not. Additionally, for the male speaker, the average F0s for the mid vowels [e], [ɛ], [o], and [ɔ] are higher than that of [i] for all three tones. The high vowel [i], along with the low vowel [a], record the lowest F0 for all three tones. For the female speaker, [i], along with [a], [ɛ], and [ɔ], record the lowest F0 for the High tone. However, in the only instance in the whole data, [a] records higher F0 than [ɛ], reflecting the finding by Hombert (1977) for the female speaker. For the Mid tone, [o] records the highest F0 while [a], [ɛ], and [i] record the lowest. The Low tone data for the female speaker indicate that

[i] has a lower F0 than all the other vowels except [a], also reflecting the observation by Hombert (1977) for the Low tone for her female subject. Meanwhile, some of the mixed results observed in the present study regarding tonal gradient mirror what was observed in Ibibio by Connell (2002), where the mid vowels [e], [o], and [ɔ] produced higher F0 than [i], for higher tones. Connell (2002) also found that [ɔ] occasionally showed lower F0 than [a], and a mid-vowel recorded a higher F0 than the high vowels [i] and [u]. Perhaps, the lack of conclusive evidence to support the presence of a tonal gradient in relation to vowel quality in Yorùbá is due to the language's larger tone inventory or the nature of the tone system, as suggested by Connell (2002).

The results are generally consistent with Hombert (1977) and Connell (2002), both of which suggest that there is no tonal gradient with respect to vowel heights. Thus, Alo (1990) remains the only study to clearly suggest that there is a relationship between IF0 and tone levels. The finding by Alo (1990) may be due to the limited number of vowels included in the study. The fact that the study investigated only the high and the low vowels - [u] and [i] versus [a] - clearly explains the results. A selective look at the data for the high vowels versus the low vowel in the present study still does not indicate such a conclusive result as suggested by Alo (1990). The data indicate that [u] is the only high vowel to consistently record significantly higher F0 than [a] for both the male and the female speakers across the three tones. The vowel [i], on the other hand, in most cases, only records slightly higher F0 than [a] (about 4 – 5Hz difference).

The average low F0 for the [i] vowel for the two speakers may be attributed to several articulatory and phonatory factors. Physiological studies indicate that different mechanisms are used for F0 control in a high as opposed to a low F0 range. Considering the statement that "IF0 will occur naturally except where phonological concerns are overriding" (Connell, 2002), it may be the case that the male speaker in this study typically has a tendency for a lower F0. It may also be the case that he has a more reduced accent (Honda and Baer, 1981) than the female speaker. These results may also suggest the involvement of other extra-linguistic factors, such as the loudness of the speech as well as the time of the day (Ogden, 2009). The recordings were made very early in the morning when the home was relatively quiet. The male speaker in this study has a normally low-pitched voice in his manner of speaking. Thus, the vocal tract shape, the larynx, and the habitual way of speaking of the speaker may have contributed to the wide variation. Moreover, as suggested by Honda and Baer (1981), it is possible that the male speaker exhibits small movements of the hyoid and epiglottis, thereby resulting in small F0 differences.

5. Conclusion

The present study investigates whether tones in tone languages constrain or control Intrinsic Fundamental Frequency (IF0). The study finds the presence of IF0 in Yorùbá. The tonal gradient is also attested, and F0 is found to be high for the High tones and reduced for Low tones. While the high vowel [u] tends to follow the prediction regarding tonal gradient, especially in comparison with the low vowel [a], the other high vowel [i] does not. In most instances, the mid vowels recorded higher F0 than the high vowel [i], thereby challenging the prediction regarding the tonal gradient effect. Thus, while the present study, in line with previous studies, highlights some differences in average F0 between the vowels, the results do not yield conclusive evidence to support the existence of the IF0 gradient with respect to vowel heights in Yorùbá. This may be due to the language's larger tone inventory or the nature of its tone system. It may also be due to some factors not captured by the study or due to the limited data available for the study. Perhaps a study involving more participants that yield more data for analysis will provide more evidence regarding tonal gradient. These are some of the factors to be duly considered in future research on IF0 studies in tone languages.

Funding: This research received no external funding.

Acknowledgments: I would like to thank the two native Yorùbá-speaking participants for providing the recorded stimuli. Many thanks also go to Professor Benjamin V. Tucker for his valuable feedback.

Conflicts of Interest: The author declares no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Alo, P. O. (1990). Interaction between segments and tone in Tsabe. *Pholia*, 5, 7-15
- [2] Bakare, C.A. (1995). Discrimination and identification of Yorùbá tones: perception experiments and acoustic analysis. In Owolabi, K. (ed.), *Language in Nigeria*. Ibadan: Group Publishers.
- [3] Beranek, L. L. (1949). *Acoustic Measurements*, New York: Wiley. Used online at the site: <<http://www2.ling.su.se/staff/hartmut/umrechnung.htm/>>
- [4] Boersma, P. & Weenink, D. (2015). Praat: Doing phonetics by computer [Computer program], Version 5.3.04. Retrieved 06 February 2015 from <<http://www.praat.org/>>
- [5] Connell, B. (2002). Tone languages and the universality of intrinsic F0: evidence from Africa. *Journal of Phonetics*, 30(1), 101-129.

-
- [6] Hayward, K, Watkins J. & Oyetade A. (2003). The phonetic interpretation of register: Evidence from Yorùbá. In J. Local, R Ogden, & R. Temple, *Phonetic Interpretation. Papers In Laboratory Phonology VI*. Cambridge: Cambridge University Press. Pp. 305–321.
- [7] Hombert, J-M. (1977). Consonant types, vowel height, and tone in Yorùbá, *Studies in African linguistics*, 8:2, 173-190.
- [8] Hombert, J.-M. & Ladefoged, P. (1976). Effect of aspiration on the fundamental frequency of the following vowel. *Journal of the Acoustical Society of America* 59(A) 572.
- [9] Honda, K. and Baer, T. (1981). Interaction between pitch control and articulatory mechanisms," presented at the 102nd meeting of the Acoustical Society of America, Miami Beach, FL, 1-4 Dec. 1981, *Journal of the Acoustical Society of America Suppl. 1 70, S12 (1981)*
- [10] Laniran, Y. (1992). Intonation in a tone language: the phonetic implementation of tone in Yorùbá. *Unpublished Ph.D. dissertation, Cornell University*.
- [11] Ogden, R. (2009). *An Introduction to English phonetics*. Edinburgh University Press. Edinburgh.
- [12] Reinholt-Petersen, N. (1978). Intrinsic fundamental frequency of Danish vowels. *Journal of Phonetics*, 6, 177- 189.
- [13] Whalen, D.H & Levitt, A.G. (1995). The universality of intrinsic F0 of vowels. *Journal of phonetics*, 23, 349-366.