The nature of Yoruba intonation: a new Experimental Study

Eunice Fajobi*

Abstract
In spite of active research work on the prosodic features of world languages in the last few decades, tone\(^1\) is the only prosodic feature in Yoruba that has received considerable attention in the field of linguistic research. Intonation\(^2\), on the other hand, has suffered a near-neglect. All there is to be flagged up in the existing literature on the intonational system of the language are mutual effects of tones on each other. But there is more to Yoruba intonation than such effects. Recent approaches do not facilitate any comparison of Yoruba intonation with that of non-tone languages. The present investigation seeks to give a more sophisticated account of Yoruba intonation.

Drawing on universal traits of intonation across languages, I report, in this paper, some of the new findings of a new elicitation study on the nature of Yoruba intonation. Acoustic analysis of the Fo patterns in some spoken Yoruba sentences in the following categories: declarative and three types of interrogative tested on four speakers of standard Yoruba, shows that Yoruba has an intonation superimposed over and above the intricacies of tone implementation reported by Carnochan (1964), Connell and Ladd (1990) and Laniran (1992). The fact that Yoruba is a tone language makes it hard to get beyond the tones to an intonation describable in such terms as head-type, nucleus-position, etc. (such as has conventionally been used to describe/analyze other languages including English). But in this study I demonstrate that Yoruba does indeed exhibit some recurrent traits of intonation such as overall fall (declination) for statements and overall rise for some types of questions in addition to patterns of lexical tones.

1. Introduction
In spite of significant advances made by researchers on the prosodic features of world languages, intonation in Yoruba is yet to receive its due share in linguistic research. Consequently, while tone has been well researched and well documented (see, for example, Siertsema 1959; Ogunbowale 1970; Courtenay 1971; Akinlabi 1985; Hyman & Schuh 1974; La Velle 1974; Hombert 1974, 1977, 1978; Awobuluyi 1978; Adewole 1987\(^3\); Oyetade 1987), the nature of the specific tunes that may characterize Yoruba as having a distinct system of intonation in addition to its lexically distinctive three tonemes remains relatively under-researched and inadequately-theorized. The reasons are not hard to discover. First, the semantic role of Yoruba tone is no doubt a crucial factor in the analysis of the phonetic and phonological features of the language. This role has, unfortunately, distracted linguists’ attention from the second major suprasegmental feature, intonation. The second and the more influential factor is the belief of some linguists (e.g. Abercrombie 1967:104; Bae 1998:38; Roca and Johnson 1999:394; and Roach 2000:162) that the world’s languages can be classified into tone and intonation languages. Asserting his view, Abercrombie says, “In every language the function of speech melody is predominantly either of one kind or the other, so that the languages of the world can be divided into two classes, intonation languages and tone languages.” This assertion,

---

* Linguistics and English Language, University of Sussex, Brighton BN1 9QN
eunicef@sussex.ac.uk

1 Tone in this context refers to the pitch of voice on individual syllable. In Yoruba, as in many African tone languages, tone is lexically distinctive. It is also phonemic/contrastive in that it behaves more like the English sounds.

2 Intonation here refers to the variation in the pitch of voice to vary meaning in utterances; “a general term for the fundamental frequency pattern of a stretch of speech.

3 Adewole, (1987) provides a list of works on Yoruba over some decades.
though non-absolute, leads to no other inference but a belief that languages can be classified as tone or intonation languages.

In the second half of the 20th century, however, native speaker intuition as well as perceptual and empirical studies are beginning to prove this notion wrong (see Pike 1948; Chang 1958; Haugen & Joos 1958; Gleason 1961; Carnochan 1964; Martinet 1964; Hombert 1974; Awobuluyi 1978; Bolinger 1978; Miller and Tench 1980, 1982; Mountford 1983; Cruttenden 1986; Lindau 1986; Ekundayo et. al 1988; Atoye 1989, 1999; Connell and Ladd 1990; Laniran 1992; Tench 1996; Atoyebi 1998). In his 1989 study, Atoye asserts that “tone and intonation do not constitute appropriate classificatory criteria for human languages as they are neither phonetic nor phonological cognates”.

Stress is the structurally valid suprasegmental feature with which to compare and contrast tone, not intonation (Atoye 1989; Fajobi 1998). In the same vein, Tench (1996:6) regards such classification of human language into tone and intonation languages as “a complete misunderstanding about the nature of intonation”. He maintains that all languages, tone or non-tone, use pitch variations for intonation purposes in their utterances and discourse. And, as also suggested by Ekundayo et al. (1988:70) for Yoruba, Tench argues that there is a tendency for a tone language to put intonation to a limited use, when compared to the use in a stress language for example. Similarly, Cruttenden (1986:10) argues that:

Tone and intonation are not completely mutually exclusive in languages. Languages with tonal contrasts may nevertheless make use of a limited amount of superimposed intonation. Such superimposed intonation may be manifested in four different ways: (i) the pitch level of the whole utterance may be raised or lowered; (ii) there will usually be downdrift in the absolute value of tones but downdrift may be suspended; (iii) the range of pitch used may be narrower or wider; (iv) the final tone of the utterance may be modified in various ways.

In this study, I show that existing studies on Yoruba intonation, such as Carnochan (1964), Connell and Ladd (1990), and Laniran (1992) do not portray Yoruba as having a system of intonation as does Hausa for example (see Miller and Tench 1980, 1982; Lindau 1986). Carnochan, Connell and Ladd, and Laniran, no doubt, provide brilliant insights into the internal structure of some utterances in Yoruba as a tone language. As a matter of fact, they offer, through the phenomenon of downdrift/downstep/declination, or upstep in some cases, a thorough exposition of the tonal interaction in some Yoruba utterances, which are, interestingly, further confirmed in the results of my study. But I argue, however, that there is more to intonation in Yoruba than the effects of tone interaction that the works report. In fact, as I demonstrate in this study, the phenomenon of downdrift/downstep/declination can be effectively blocked by some superimposed features of intonation in some

---

4 This quote is from Atoye’s 1999 work. The original claim is made in Atoye, (1989:1)
5 Fajobi, 1998 in her study of word stress placement problems of teachers of English in secondary and tertiary institutions in Nigeria stresses (after Atoye) the high correlation between the three pitch levels of Yoruba tonemes (H, M, and L) and the English primary, secondary and tertiary stresses.
6 Contrary to the claim by Connell and Ladd (1990:27), the final rises for the all high and all mid tones in the renditions of the informants in my study for the declarative sentences are absolutely negligible. The final drop in F0 for the all L tone statement is also more pronounced with the female subjects than with the males (cf. Fig. 4.4).
utterance types in Yoruba. Further, though the syllable by syllable (or tone by tone) examination of utterances approach, which Roach (2000:162) describes as common to tone language researchers, yields parallel results for these researchers (Connell and Ladd 1990:5), the approach does not adequately capture the nature of intonation in the language. Consequently, there are yet to emerge any specific patterns of intonation in Yoruba that can be compared with the intonation system of other well-established tone or non-tone languages. In this study, I analyse the final Fo trajectory of some Yoruba declarative and interrogative sentences to prove that Yoruba, in spite of its lexically distinctive tones, uses intonation. I will show that as a matter of fact Yoruba uses different intonation for different utterance types. In addition, I will follow Cruttenden’s (1986:10) and Tench’s (1996:6) positions to show that though tone in Yoruba may be functionally semantic, it is neither synonymous with, nor replaces the linguistic or the functional roles of intonation in the language.

It should be noted from the outset that this study is by no means an attempt to contradict the reports relating to the phenomena common to the phonetic implementation of tone in Yoruba. Rather, it complements existing findings through a more sophisticated approach to establish a well defined intonation system for the language. Further, I do not claim an exhaustive account of Yoruba intonation in this study. It is limited in scope to declarative and some interrogative utterances only. It is believed, however, that the findings will, in addition to attracting interest in further studies on the intonation of attitudes, vocatives, clausal relationships, focus, and general discourse in Yoruba, make the system readily available for comparison with the intonation system of other languages, tone or non-tone. Given the increased interest of linguists in bilingual and multilingual studies in recent times, a study such as the present one becomes crucial, and perhaps an invaluable asset.

This research study aims, then, at:
1) Establishing the use of overall intonational tunes in Yoruba superimposed on its well established tonal system;
2) Providing a more radical approach to analysing and to establishing the patterns of intonation found in the language; and
3) Providing the basis for constructing and for showing the interrelatedness of tone and intonation in Yoruba.

2. Yoruba Prosody: Reflecting on past works
2.1. Socio-cultural setting

There is a significant population of Yoruba around the globe. Users of the language have been said to be up to 100 million. This include the Yoruba people of the south-western part of Nigeria and their remnants found in many parts of the world such as in Cuba, Brazil, and in other parts of South America. Yoruba is the predominant language of the people in the south-western part of Nigeria. It belongs to the Kwa language subgroup of the Niger-Congo family (see Williamson 1979). It is a tone language. About 37% of the Nigerian populations who use Yoruba as their mother tongue are located in Lagos, Oyo, Ogun, Oshun, Ondo, Ekiti, and in some

7 Connell and Ladd attested to Laniran’s results as being parallel to theirs. The works were carried out simultaneously, but at different locations. Carnochan’s (1964) results are equally similar in almost every respect to the results in the above works.
8 www.nmbacom.com/21CenturyYoruba/21stCenturyYoruba.htm/ Most of the information in this paragraph was downloaded on 25 January, 2004 from this web address, and in the next note.
9 Williamson is a rich source for the tone languages of the Niger-Congo family. Her 1979 article, cited above is a study on some southern Nigerian tone languages.
parts of Kwara and Kogi states. It is also spoken in Benin Republic. There are as many as seventeen dialects in Yoruba (Herault 1981). The Oyo dialect, which is the well-acclaimed standard variety used in schools and in Yoruba literature, is the variety I shall allude to in this study.

2.2. Yoruba Tonology

Tone in Yoruba, and indeed in all African tone languages (Pike 1948:3), including the Bantu languages (Cole 1955:53-55), is semantically significant, i.e. it serves to distinguish between words having different meanings, but which are otherwise phonetically identical. In descriptive, theoretical and acoustical terms, tone in Yoruba has been widely and extensively studied (see Ward 1952; Olmstead 1953; Siertsema 1959; Carnochan 1964; Bamgbose 1966; Ogunbowale 1970; Courtenay 1971; Welmers 1973; Hyman & Schuh 1974; La Velle 1974; Hombert 1974, 1977, 1978; Awobuluyi 1978; Akinlabi 1985; Adewole 1987; Oyetade 1987; Connell and Ladd 1990; Laniran 1992). Phonologically, Yoruba has three contrastive tones—High (H), Mid (M), and Low (L). The syllable is the domain of tone: every syllable in the language is usually associated with a tone. Consequently, syllable tones are of semantic importance for differentiating between words having identical sequences of segmental phonemes. I illustrate the tonemes (contrastive tones) with the following examples:

1. ri (H) see
2. ri (L) sink (also rot)
3. ri (M) collapse
4. ba (H) catch up with
5. ba (L) rest on (also to ferment)
6. ba (M) hide

In the system’s standard orthography, every syllable is tone-marked. But while the M tone-bearing syllable is conventionally marked implicitly, i.e., without any diacritical mark on such syllables, the H and L tone syllables are marked (/) and (\) respectively. In the past, some Yoruba linguists attempted marking M tone syllable with a macron (i.e. a diacritic ¯ over the letter as in <ā>), but that practice has long faded out.

The association of every syllable in Yoruba with a tone allows a theoretical claim that tone configuration in Yoruba is complex and free; free in the sense that any of the three tones may occur on the syllable of any phonological shape but complex in that there are semantically conditioned tonotactic restrictions. Although it is possible to achieve special effects like those illustrated in (7) through (9) which might imply complete freedom of tonal choice, there are nevertheless constraints on the sequence of tones that can occur on polysyllabic words. For instance, not all of the nine different potential tone configurations possible in Yoruba are actual word-shapes in the language (Atoye 1999:15). In the following examples, only (10) through (14) are actual Yoruba words; (15) through (18) are not. One clear inference that can be drawn from all these is that phonological construction in Yoruba may be guided by some complex tone rules. This is beyond the scope of the present study.

10 <http://www.isp.msu.edu/AfrLang/Yoruba_root.html>
11 Examples are mine.
7. Won tun gbe tuwo wa (All H tone)
   ‘They brought tuwo again’
8. Omo won ni e lo fi s’oko (All M tone)
   ‘It is their son that you marry’
9. Ewu ona Ara o tan (All L tone)
   ‘The colour of the garments on the way to Ara is dull’

10. igba (MH) — calabash
11. igba (LL) — time
12. igba (MM) — 200
13. igba (LH) — garden egg
14. igba (ML) — rope used by wine-tappers
15. igba (LM) — ?*
16. igba (HH) — ?*
17. igba (HM) — ?*
18. igba (HL) — ?*

Following from the above, the understanding of tone is no doubt crucial to the understanding of the language. This, as earlier observed, may be one vital factor that leads many linguists to assume that all there is to Yoruba prosody is lexical tone, and the reason so much attention has been paid to it to the neglect of intonation.

2.3. Yoruba in the typology of tone

Pike (1948) classifies tone languages into two broad categories – the register tone system, and the contour tone system. By Pike’s classificatory criteria, Yoruba belongs to the register tone system, a system characterized by level tones against the pitch contour tones used in certain Asian and American-Indian languages. Within the register tone system, Welmers (1973:80) recognizes two types of tone languages: terrace tone and discrete level tone languages. It is common practice among researchers in African tone languages, following Welmers’ definition of the terminology, to draw a dichotomy between discrete level and terraced level tone languages (cf. Connell and Ladd, 1990:3; Laniran, 1992:12). In a discrete level tone system, the pitch ranges of level tones are distinct. They do not overlap. In any given utterance of H, M and L tones, the Fo value of every H must be higher than that of every M, and that of M must in turn be higher than the value of every L (Welmers 1973:80). In a terraced level tone system, “each high tone in a sequence is lower in absolute pitch than the preceding one” (Ekundayo et al 1988:69 after Welmers 1973:82). In support of Connell and Ladd (1990:4), I argue that “a dichotomy between these two types” is indeed, “logically unnecessary” as well as empirically unsupported because the two can coexist in any one language as demonstrated by findings in this study (cf. Figures 4.2, 4.3, 4.13, and 4.14). What I observe in this present study is that the overall pattern is always dependent on tone sequence and the length of utterance on the one hand, and on the explicitly marked speech act performed by the utterance on the other hand. The speech act seems to exert more influence. Miller and Tench (1980:52ff) report this phenomenon for yes/no questions in Hausa. It is also confirmed

---

12 Examples 7 - 8 are taken from Connell and Ladd 1990, p.8. There is nothing to suggest that the sentences are stereotypical.
in this study for the Yoruba yes/no question with particle in the final position (e.g. Fo contour for “...bi?” question in Figure 4.14). While terracing or downstepping is clearly evident in the all H tone statement (Figure 4.3), the pattern towards the end of the all H question-word question (Figure 4.14) shows a reverse of “normal” pattern, which Connell and Ladd have described, in their study, as “hogback ridge pattern” (see p. 4). The only argument that can be possibly proffered for this occurrence is the effect exerted on the utterance by the final syllable being a question word — an instance of a superimposed intonational effect. This explanation becomes tenable in the absence of such Fo increase in the final position for the all M or all H declarative utterances. Furthermore, when decontextualized for pedagogic reasons for instance, Yoruba tones will manifest discreteness. In the light of the above, I suggest that Yoruba be simply described as a register tone language.

2.4. Yoruba Intonation- Past works
2.4.1. Downdrift, Downstep, Declination

As mentioned in Section 1, the notion that a tone language could use intonation is yet to gain acceptance of some linguists. Pointers to this observation are the 20th Century authors (e.g. Bae 1998:38; Roca and Johnson 1999:394; Roach 2000:162) who still describe world languages as tone and intonation languages; this stance, no doubt, distracts attention from the study of tone language intonation. It is no wonder then that there is no distinction yet between the theory of tone and intonation in descriptions available for most African tone languages. Analysis is always focused on the mutual effects of tones on each other, often with reference to the phenomenon of downdrift; but downdrift is peculiar more to the rules guiding tonal interaction than to the rules of intonation per se. As an instance, while the phenomenon works for the description of statement intonation, where, obviously, tonal interaction plays crucial roles, it does not give a clear understanding of the intonation of tone language questions.

Connell and Ladd (1990) and Laniran (1992), who all did instrumental studies of tone realization in Yoruba, are best described as offshoots of works on Yoruba tone such as Courtenay (1971), La Velle (1974) and Hombert (1974). Essentially, their findings, which portray the assimilatory and intonational nature of downdrift, which Hombert (1974:171) describes as generally assumed by tonologists, replicate a detailed account of what Pike (1945:60, 1948:60ff.) describes as the necessary internal structure of intonation; in this case the underlying local rules of tone over which the speaker more or less has no control. Reports common to these works for Yoruba include among others that:

- There is a H raising triggered by a following L in a sequence H-L in Yoruba. Lindau (1986) and Mountford (1983) report the same for Hausa and Bambara respectively. The low tone in the sequence may be raised or realized as mid (Hombert 1974:171).
- Hs following Ls are lowered (also reported by Lindau 1986:759 for Hausa)
- Automatic lowering of both H and M after L (also reported by Courtenay 1971:240, when explaining the phenomenon of terracing in Yoruba).

---

13 Downdrift, downstep and declination manifest the same phenomenon. They are used to refer to the progressive lowering of the pitch of voice when speaking, i.e. the gradual decline of Fo (pitch) during the course of utterance (Ladd, [Web article]). Connell and Ladd (p.2) observe that while declination is used more for European languages, downdrift is used for African tone languages. Hombert, (1974:169) describes the phenomena as universal for establishing a theory of tone.
- Sequences of like tones exhibit a gradual fall in Fo which is dependent on tone types. In the process, L tones show more declination (i.e. more lowering) than M tones and H tones show less declination than M tones. All of these are confirmed in the findings of the present study as discussed in Section 4.1.

In addition to the above, Connell and Ladd (1990:15) made a brief observation that questions in Yoruba are likely to exhibit wider pitch range than statements. Their observation has been confirmed in a broader sense in my experiment. In a related work, Carnochan (1964:398) justifies Yoruba as a language that uses different pitch range for statements (nr-normal register) and questions (hr-high register). In his analysis, Carnochan assumes intonation as “the sequence of the tones of the syllable corresponding to the pronunciation of the example”. Indeed, all attention in the analysis is paid only to the corresponding syllables and tones in the two-syllable-utterance examples used, and it becomes difficult to distinguish between the characteristics of statement and questions except the labels ‘nr’ and ‘hr’, which, at best, portrays Yoruba as having only one question type. On the other hand, Laniran (1992), whose study centres on Yoruba declarative utterances only, makes no mention of questioning at all. The general inference to be drawn from the above findings is that Yoruba lacks any specific pattern of intonation aside of declination/downdrift phenomenon common to declarative statements, or that all there is to be talked about in the prosody of the language is the mutual effects of the lexical tones on each other. The account is more like the analyst of a stress language, who concerns himself or herself with the analysis of the rules governing the co-occurrence of the individual stressed and unstressed syllables in utterances without any apparent reference to the overall intonational tune superimposed on the utterances. Based on this notion, the present study will explore the following research questions:

1a) Does Yoruba as a tone language use intonation at all?
1b) if so is intonation in Yoruba a mere occurrence of lexical tones in sequence or are any other features of intonation added?
2) If Yoruba uses intonation, to what extent does the system/pattern conform to the universality of intonation in the majority of world languages – i.e. falling for statements, and rising for questions? (Bolinger, 1978)
3) Can any difference be established between the intonation of statements and questions?

3. Method
3.1. Theoretical framework

In addition to establishing whether or not Yoruba uses intonation in addition to the lexical tones, this work aims specifically at defining for Yoruba an intonation pattern that makes it readily available for comparison with the intonation patterns of other world languages, tone or non-tone. To this end, I follow Lindau’s (1986) model for the intonation of Hausa, another major Nigerian language with two tones against the three used by Yoruba. Lindau’s is a modified version of that of Gårding and Bruce (1981) for describing the intonation of Swedish, a pitch accent language. The model, described by Lindau (1986:757) as layered, provides a framework for separating the rules for tone and intonation. In the model:

---

The pitch contour is determined by a combination of global and local factors. The global, overall intonational slope is described using an abstract ‘grid’ of two (near) parallel lines, a top line, and a base line. The Fo contour is determined by the placement of sequences of linguistically significant local high and low tones as targets inside the grid. The final Fo trajectory is generated by rules of interpolation between the targets.

Though Mountford (1983), in his description of Bambara\textsuperscript{15} declarative sentence intonation, reports the Gårding and Bruce model as perfectly suitable for his purpose, the simple reason that a three level tone language like Yoruba may not be easily represented inside two parallel grid lines has necessitated a further modification of the intonational model adopted for the present study. In this new model, Lindau’s model is combined with the model used by Marcé (1998) in her study of how Spanish subjects acquire English intonation. In Marcé’s model, mean values of corresponding stressed and unstressed syllables were exported into the Excel spreadsheet to generate an ‘intonation graph’ for individual utterances as rendered by the informants. In the final intonational curve, she was able to determine the nucleus positions for the informants’ declarative, interrogative, imperative, and exclamative utterances. Inspiration for combining this approach stems from two main sources: (1) Mountford’s (1983: xxii) submission that ‘register tone languages and stress languages can be quite similar in their phonetic realization of intonation.’, and (2) Pierrehumbert’s (1980) highly influential autosegmental theoretical approach in which she uses a theory of tone to describe a stress language intonation.

For the present study, the Fo trajectory for Yoruba declarative and interrogative utterances was determined as follows:

1. The fundamental frequency values of the linguistically significant local high, mid and low tones were extracted from the informants’ sentence by sentence renditions using the pitch extraction software, PRAAT (see Sections 3.2 and 4);
2. The mean Fo values were exported into the Microsoft Excel spreadsheet, which converted the local targets into a recoverable global intonation curve at the application of superimposed functions;
3. Comparison was made between the Fo contours generated, using statistical methods. For the statistical analysis, SPSS was used. The anchor point (AP) was paired against the end point (EP) in a 2-tailed paired sample tests. The anchor point represents the first highest point (i.e. the tone with the highest Fo value) in the intonation graph while the end point is taken as the lowest Fo value in the curve. But the EP tone must be similar to the AP tone so that if AP is a high tone, EP too must be a high tone. In each case, the end point corresponds to the utterance-final syllable.

One advantage of exporting the Fo values into the Excel spreadsheet, rather than use the raw pitch curve generated in PRAAT, is that actual and accurate mean values are easily represented on the graph. This in turn makes the determination of various points such as the anchor points and the end points needed for statistical analysis readily available.

### 3.2 Data Collection

Data For this study was gathered by elicitation method in 2002. A Sharp portable MiniDisc recorder, model MD-MT88H, was used for the recording of all the

\textsuperscript{15} Bambara is a two-toneme (high and low) language similar to Hausa (see Mountford, 1983).
sentences. Before being fed into the pitch extraction software, PRAAT, for further and detailed analysis, the recorded sentences were run unto a Macintosh computer (Apple Mac G4 running Mac OS X) on a 16-bit mono sound file at a sample rate of 44100 Hertz at time step 0.01 seconds. These utterances were later sliced up into smaller manageable files, using the Sound Studio application. Every file was converted into AIFF (Audio Interchange Formant File) sound in the object window of the analyzer through the Write-menu of the programmer. For the extraction of Fo values in this study, the pitch range was set at 75 - 300 Hz and 75 - 500 Hz for males and females respectively, following Ladefoged (2001:165), and Kent and Read (2002:191). Because the Fo of a male voice in speech does not normally exceed 300 Hz, and that of a female may not be higher than 500 Hz, the default setting in PRAAT, which was 75 - 600 Hz, was adjusted as discussed above. The adjustment does not affect the output of the software in any way.

For the segmentation of the utterances into the tone-bearing units, i.e. syllables (after Goldsmith 1976; Clements and Ford 1979; Odden 1995), fundamental frequencies were traced and extracted from the spectrograms of the utterances, using several superimposed commands in PRAAT window. The software also made possible a recovery of values for the Fo, duration and intensity of each syllable. This was recorded manually. I should admit here that the acoustical analysis stage proved the most difficult and time consuming encounter in the study. It also accounts for the limited number of utterances used for the study. The analysis was completed statistically.

I deliberately avoided conversational speech for a number of reasons, which include empirical shortcomings that might arise in the event that the test items of the required type are not obtained. For this reason, I decided to go for elicitation method. Before each recording, the informants were given ample time to read through the questionnaire as many times as they wanted until they were sure they could read fluently like they would in natural speech. Where they made mistakes, they were allowed to repeat the reading. This way, the signals generated for the study are empirically representative of natural speech in the sense that they satisfy the speaker’s own criteria for this, and the results can be generalized to form a model of what obtains in natural speech. Any sentence read with different tones for instance, was thrown out of the analysis. The questionnaire consists of first, the test items in Yoruba, and second, Yoruba sentences with their English glosses16. In subsequent analysis, however, there was no significant difference in the two outputs, so I chose the pure Yoruba utterance outputs as my data.

3.3 Sentence Materials

Eleven Yoruba sentences, properly tone-marked in accordance with standard spelling-conventions, were used to elicit the data for this study. They were constructed to establish the pitch patterns for Yoruba declarative and three types of interrogative. Example sentences 19 – 23 were used to test the tune pattern for Yoruba declarative; 24 for yes/no question without particle; 25 – 27 for question-word question (or yes/no questions with particles); 28 and 29 for Wh-clause questions17.

---

16 The purpose of eliciting the English glosses is explained appropriately elsewhere in my dissertation. Nevertheless, the tokens thus generated have proved helpful in “validating” the data for this study.

17 Wh-clause question as used in this study refers to a subtype of interrogative clause usually “marked by the presence of one or more of the interrogative words who, whom, which, whose, what, where, when, why, how…” (Huddleston, 1984:365-6).
3.3.1. Example Sentences for Declarative
19. Omo naa ti de. (MM HL M H)
   (The child has arrived.)
20. Ara won ni Opara. (MM M M MMM)
    (Opara is one of them.)
21. Sikira lo mowo. (HHH H HH)
    (Sikira took the money.)
22. Tete ategunmo nijangbon opele. (LLL LLL LLL LLL)
    (Tete and egunmo are the problems of opele.)
23. Nkan ti o a so ni pato ni pe Tejumola fe segbeyawo.
    (What he is saying in essence is that Tejumola wants to get married.)

3.3.2. Example Sentence for yes/no question without particle
24. Omo naa ti de?
    (The child has arrived?)

3.3.3 Example Sentences for question-word question or yes/no question with particle
25. Se omo naa ti de?
    (Has the child arrived?)
26. Omo naa ti de bi?
    (Has the child arrived?)
27. Abi omo naa ti de?
    (Could it be that the child has arrived?)

3.3.4 Example sentences for Wh-clause questions
28. Bawo ni e se mo pe omo naa ti de?
    (How did you know that the child has arrived?)
29. Kini idi ti omo naa fi de?
    (What is the reason for the child’s arrival?)

3.4 Subjects
Four of the 26 informants, who participated in the experiment conducted for my ongoing doctoral research, constitute the subjects for this study. They are final year BA Yoruba students of the Obafemi Awolowo University, Ile-Ife, Nigeria. The

---

18 “Tete” and “egunmo” are names of some vegetables in Yorubaland; “opele” is one of the instruments used for “Ifa” divination.
19 Only one sentence example is used here because no other dataset could be formed from the core sentence “omo naa ti de” (see 3.3)
20 Wh-clause question as used in this study refers to a subtype of interrogative clause usually “marked by the presence of one or more of the interrogative words who, whom, which, whose, what, where, when, why, how…” (Huddleston, 1984:365-6).
21 Going by Huddleston’s definition of Wh-clause questions and their introducers, sentence 29 can similarly be translated as “Why did the child arrive?”
informants, two males and two females, within the age range of 21 and 25 years, are speakers of standard Yoruba, who had, in previous years, taken courses in general phonetics and phonology in addition to studying the phonetics and phonology of Yoruba. They were randomly selected from among those with an average GPA of second class upper. For this work, they are identified as JIDM, COMF, OPEM and TOLF. The last letter in each name indicates informant’s sex.

4. Analysis, Results, and Discussion

In consonance with my aim of extracting the values for fundamental frequency, intensity and duration to analyze intonation in this study (following Zee, 1978, and Marcé, 1998), the corpus generated was fed into the pitch extraction software, PRAAT, designed and maintained by Paul Boersma and David Weenink of the University of Amsterdam in The Netherlands. The version used for the present study is Version 4.1.3. The software enables a great variety of acoustic analysis. For instance, it can be used to analyze, synthesize, and manipulate speech sounds. At the application of appropriate commands in the object window of the programmer, varieties of acoustic signals including spectrogram, formant, pitch and amplitude traces are displayed. At a further application of superimposed functions, fundamental frequency values of speech sounds fed into the object window as well as the values for the amplitude and the duration are shown in the “Info” menu of the programmer. Getting the pitch contour of a given utterance is quite easy with the programmer. One other advantage of this software package is that, users are able to add grid lines and texts to the drawings of extracted speech sounds (see figure 4.6). In addition, values can be added on the y- and x-axis of the pitch or amplitude contours. Furthermore, superimposed pitch contours can be drawn in the PRAAT picture window for an easy comparison (see Figures 4.13 and 4.14 for example). Apart from requiring a mastery of various key-applications, commitment, and a high level of concentration and time devotion, PRAAT aids accurate segmentation of utterances into syllables/tones or phonemes through the spectrographic display of speech files.

Perhaps it is worth mentioning at this juncture that the values represented on the graphs are mean values of the informants’ syllable by syllable productions. This method helps to minimize variations in the range of the informants’ pitch of voice, which might have had some negative effects on the results of the experiment if not controlled. This also explains why the ‘raw pitch contours’ may vary slightly from the way they are represented on the graph. Over and above all this, the emerging graph curves for the sequence of tones in the utterances provides a way of representing the informants’ Fo contours more uniformly. Comparison on an intonational plane, regarding the common characteristics that can be generalized, similarly becomes easier. This pattern of intonation reflects more accurately the way we perceive the melody, since our ears work with a logarithmic scale, not a linear one (Marcé, 1998).

4.1. Experiment 1: Testing the intonation of Yoruba declarative sentence

In this experiment, five declarative sentences consisting different tonal configurations (example sentences 19-23 [see 3.3.1]), were tested on four informants to establish the pattern of Yoruba declarative intonation. Figures 4.1 through 4.5 show the intonational graphs generated for the utterances, using the model adopted for the study. The four intonational graphs in each figure represent the global, overall intonational slopes mapped out of the linguistically significant local Fo targets exported into the Excel spreadsheet for the four informants (see the graph legend for easy identification of the informants). As evident in the figures, there is a high degree
of uniformity in the overall Fo trajectory generated by the informants for each utterance type. However, there is variation in the internal structures of the utterances.

Figure 4.1: Intonation graph for the short declarative statement with mixed tones (P = .004).

Figure 4.2: Intonation graph for all-M declarative sentence (P = .052). Informants OPEM and TOLF’s final syllables are lowered; those of COMF and JIDM are raised, but the raising is negligible.

For some of the figures, the variation tends to hinder a judgmental claim that the utterances have one feature in common that characterizes them as belonging to one category intonationally; i.e. they all exhibit a common trait of statement intonation. This is not much of a surprise because as rightly observed by Laniran (1992: 205), the rising and falling patterns that will be realized phonetically will always be dependent on the interaction between the number and the sequences of the tones in the utterance. On the other hand, however, her claim that the seeming irregularities in the pitch contours for different tonal configurations would make it impossible to posit specific pattern(s) of intonation for the three-toneme Yoruba as did Pierrehumbert (1980) and Liberman (1975, cited by Pierrehumbert, 1980)) for English, is unsupported by the findings in this present study. For example, Laniran observes declination as an intonational phenomenon in Yoruba, but she almost limits her use of the terminology to sequences of like tones only. As a result, her intonation of “mixed tone” utterances,
which she explores through the phenomena of downstep and “left-branching” upstep, is characterized by the effects of mutual interaction between the tones. Such effects, which include among others, H-Raising in HL sequence; tendency for Fo value of H to increase in a sequence LHL; long distance dependency in Fo realization between H

![Figure 4.3: Intonation graph for the all-H declarative sentence (P = .055).](image1)

![Figure 4.4: Intonation graph for the all-L declarative sentence (P = .011).](image2)

tones and L tones in adjacent environments, etc., are all confirmed by findings in my experimental study as they are in Connell and Ladd (1990), who simply reported aspects of pitch realization in Yoruba. Furthermore, Laniran’s characterization of Yoruba as “a lexical tone language with three level tones, declination (which varies according to tone) and rules of partial automatic downstep and H-Raising” (pp.12-13) gives the impression that her work is, essentially, an advanced form of the works of Yoruba tonologists such as Courtenay (1971); La Velle (1974); Hombert (1974, 1977); and Clements (1979). One then wonders if there is indeed anything like a separate intonation superimposed on Yoruba lexical tones, and whether or not the perturbations caused in sequences of tones would allow any analysis that goes beyond tonal interaction to show the language as having specific intonation pattern(s) in spite of the “problems” associated with the phonetic realization of the three lexical tones.
To the best of my understanding, every global intonation is generated from some underlying local rules, which Pike (1948:60-1) describes as essential to intonation. And while the speaker more or less has no control over the internal local rules (see also Fajobi 2003), the fact that he has “an idea of what he is about to say” (Cohen and ‘t Hart 1967:179) gives him the choice of how to say it to project the desired meaning to his listener (Fraser 1977:164; Cruttenden 1997:7; Roach 2000:155-58). This may be expressed in speech-act terms. As an instance, statement intonation is reflected in a universal connotation of falling, while rising is usually associated with some types of questions (see Bolinger, 1978). Be that as it may, I regard all the rules associated with tonal interaction as the inevitable internal local rules over which the speaker has no control. Consequently, the final Fo trajectory common to Yoruba speakers for statements and questions shall be the major concern in this exploratory study.

As explained in Section 3.3 and shown in Table 4.1 below, five sentences are used to test the nature of declarative statement intonation in Yoruba: there is a short sentence with mixed tones (only six syllables); three, specially constructed but not too long, sentences (all mid tones, all high tones and all low tones); and a fairly long sentence (twenty syllables) with mixed tones. The general motive behind constructing the declarative sentences thus is to draw comparison between their levels of declination.

Table 4.1 provides statistical results for the five statements studied in this section. As shown in the table, there appears to be no significant difference in the direction of pitch movements in the renditions of the informants. Each statement, irrespective of the tonal configuration, shows very interesting traits of declination, i.e. each statement declines significantly.
Table 4.1: 2-Tailed paired tests For Declination in Yoruba declarative sentences (S-Mix [example 19]) = short sentence with mixed tone; All-M (example 20) = all mid tone; All-H (example 21) = all high tone; All-L (example 22) = all low tone; Lg-Mix (example 23) = long sentence with mixed tone.

<table>
<thead>
<tr>
<th>Statement Type</th>
<th>S-Mix (19)</th>
<th>All-M (20)</th>
<th>All-H (21)</th>
<th>All-L (22)</th>
<th>Lg-Mix (23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>.999</td>
<td>.999</td>
<td>.993</td>
<td>.974</td>
<td>.964</td>
</tr>
<tr>
<td>Pearson's Sig.</td>
<td>.001</td>
<td>.001</td>
<td>.007</td>
<td>.026</td>
<td>.036</td>
</tr>
<tr>
<td>t-ratio</td>
<td>8.193</td>
<td>3.124</td>
<td>3.069</td>
<td>5.608</td>
<td>4.038</td>
</tr>
<tr>
<td>Sig.(2-tailed)</td>
<td>.004</td>
<td>.052</td>
<td>.055</td>
<td>.011</td>
<td>.027</td>
</tr>
<tr>
<td>N</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

It is to be noticed, however, that though the level of significance varies by tonal configuration, and the size of data constraints over-generalization, the consistency of the informants’ output and the statistical results indicate a very high probability of the findings as representing something significant about Yoruba declarative statements. For the short sentence with mixed tones (S-Mix [example 19]), the correlation coefficient between the anchor point and the end point is significantly high at 1% level (r = .999). In a paired sample test, this statement shows a statistically significant declination with a p-value of .004 (t = 8.193; N = 4). The all mid tone statement (All-M [example 20]) declines at a statistically significant level of 5% (r = .999; t = 3.124; p = .052; N = 4); all high tone statement (All-H [example 21]) declines at a statistically significant level of 10% (r = .993; t = 3.069; p = .055; N = 4); the all low tone statement (All-L [example 22]) declines statistically significantly at 1% level (r = .974; t = 5.608; p = .011; N = 4); and the long statement with mixed tones (Lg-Mix [example 23]) declines at a statistically significant level of 5% (r = .964; t = 4.038; p = .027; N = 4). The implication of all these is that all the statements, in statistical terms, exhibit a downward slope in a left to right gradient that shows that statements in Yoruba are said on a falling tone.

The short statement with mixed tones, which shows the least declination graphically (see Fig. 4.1) declines most with a p-value of .004. When compared with Lg-Mix (Fig 4.5) with a p-value of .027, the difference in the level of declination between the two provides further evidence for the claim that short sentences have steeper slopes than long ones (e.g. Ohala and Ewan 1972; Hombert 1974). Hombert (1974:171) argues that the difference is “caused by the fact that each speaker has physiological constraints as far as his pitch range is concerned and consequently can not maintain indefinitely the same interval of lowering”. Similar findings have been reported for Hausa (Lindau 1986), and for Bambara (Mountford 1983). Another interesting finding that confirms existing works on Yoruba (e.g. Connell and Ladd 1990, and Laniran 1992) is that an all-low-tone utterance declines more in absolute
pitch value than the all-mid-tone, and the all-high-tone declines least (see Figures 4.2, 4.3, and 4.4). The p-values are .011, .052, and .055 respectively. In other words, when each of the statements is said, there is a progressive lowering of the speaker’s pitch of voice over the course of the utterances unlike in Dschang, where sequences of high tones do not decline (Hombert 1974:178). At this juncture, it can be safely concluded that all things being equal, a Yoruba declarative statement is said on a falling tune. This in turn makes the language to share with most world languages the universality of statement intonation as observed by Bolinger (1978), Clark and Yallop (1990), and Ladd (1996), and in language specific works such as Chang (1958), Pierrehumbert (1980), Lindau (1986), Bae (1998), Newman (2000), Jaggar (2001), Flynn (2003). In these languages, statement intonation exhibits and maintains a falling tune.

Commenting on the universality of falling tune as a trait of declarative intonation in world languages, Bolinger (1978:471) says:

The traits of intonation shared by majority of languages, not excepting tone languages, are both formal and semantic, and cover the two main non-tonal uses of pitch variation: to form closures (descending lines, clause-final falls and non-falls) [...]. Terminals are almost universally low or falling for finality and assertion, and high or rising for the opposite, including yes-no questions [...]. Most deviations from the central tendencies can be explained in reference to those tendencies…

In his experiment on the intonation of Chengtu dialect of Szechuan in China, a tone language, Chang (1958) reports the declarative sentence as having a falling tune. Chang asserts further that, “If the sentence ends in a high tone while the rest of the sentence are low tones the high tone naturally remains higher than the low tones, but even the high tone has an inclination to fall.” i.e. intra-syllabically [p. 408]. Similarly, Lindau (1986); Miller and Tench (1980); and Jaggar (2001) observe a progressive downward sloping of the pitch of voice for normal Hausa declarative intonation. Apparently struck by a similar finding for Bambara declarative sentence intonation, Mountford (1983:185) concludes that “register tone languages and stress languages can be quite similar in their phonetic realization of intonation.” Findings in the present study show that the pattern of Yoruba intonation is similar to the pattern in English, except that the Yoruba pattern is not based on nucleus identification as is customary for English (see Section 5).

4.2 Experiment 2: Testing the intonation of questions

Yoruba has a very rich, flexible, and easy to identify system of questioning. Users might seem to rely on syntactic modification of canonical sequences by inserting extra words to ask questions, but in the language, a question may have no question word, it may be preceded or succeeded by a question word, or be placed between two interrogative particles (Ward 1952) as in the following (examples 24 - 27) respectively:

24. Omo naa ti de? (Has the child arrived?)
25. Se omo naa ti de? (Has the child arrived?)

22 Emphasis is mine.
26. Omo naa ti de bi? (Has the child arrived?)
27. Abi omo naa ti de (ni)?
   (Could it be that the child has (actually) arrived?)

In other instances, questioning may be introduced by a Wh-clause as in:

28. Bawo ni e se mo pe omo naa ti de?
   (How did you know that the child has arrived?).
29. Kini idi ti omo naa fi de?
   (What is the reason for the child’s arrival?).

Context too may require lexical substitution, which may or may not infringe on the tone of the word being substituted. In (29) for instance, fi is substituted for ti (before “de”) in other datasets, because (30) is ungrammatical. It is to be noticed however that this substitution has no effect on tone because both fi and ti have similar tones (i.e. mid).

30. *Kini idi ti omo naa ti de?

Apart from the zero-particle question (e.g. 19), which may pose no problem for a perceptual analyst because of the relatively higher pitch with which it is said compared to its statement counterpart which declines automatically by virtue of being a statement, the presence of question word(s) coupled with the lexical tones are salient enough to make analysts assume that these utterances do not have any intonation superimposed on them beyond sequences of lexical tones, which would, essentially, make the realizations of the canonical sequences in sentences 25 - 29 similar in every respect to the realization of statements. Conversely, acoustic analysis in the present study has enabled an interesting discovery that, when grouped together into categories as discussed below, these questions exhibit recurrent traits of varying degrees of superimposed intonation. The three types explored in this study include: (1) Yes/No question (or zero-particle question), (2) Question-word question, and (3) Wh-clause question.

4.2.1. Yes/No questions

Most world languages have been reported to mark the intonation of yes/no questions on an overall higher pitch range than for their statements (see Bolinger 1978). In Yoruba, yes/no questions are of two types: zero-particle question and question-word question. As shown in my data, the mean overall pitch in both types are uttered on higher pitch range than statements and Wh-clause questions, but the pitch is significantly higher for the zero-particle question (see Fig. 4.6). On the one hand, the question-word questions in this study are pragmatically similar to the zero-particle question in that they too can generate yes/no responses as in:

<table>
<thead>
<tr>
<th>Question-word question</th>
<th>Yes/no response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Se omo naa ti de?</td>
<td>Beeni/Rara</td>
</tr>
<tr>
<td>(Has the child arrived?)</td>
<td>(Yes/No)</td>
</tr>
</tbody>
</table>

23 No such example in this study. What I have instead is a construction without the final particle “ni”; inclusion of it may be counterproductive for my experiment.
On the other hand, they can elicit “link-answer” responses as in:

**Question-word question**  **Link-answer response**
Se omo naa ti de?                      O ti de/Ko tii\(^{24}\) de.  
(Has the child arrived?)             (He has/He hasn’t).

Figure 4.6: F0 contours for comparing JIDM’s STM and Z-PARQ utterances in PRAAT (i.e. declarative (19) and its zero-particle question (24) counterpart). The range between the two utterances is narrower at onset than utterance finally. This suggests a raising of utterance-final syllable.

Figure 4.7: Intonation graph for zero-particle question “Omo naa ti de?”

\(^{24}\) In the “link-answer response” column, we notice that vowel /i/ lengthens; the linguistic effect is to show contrast between the phrase “ti de” (has arrived) and “tii de” (has not arrived”. It functions as a negating particle.
4.2.1.1. Zero-particle question

One common characteristic feature of zero-particle questions, also referred to as “statement questions” or “declarative questions” (see Geluykens 1986:16; Cruttenden, 1997:155, respectively), is that they are structured like statements. They are grouped under yes/no question because they “ask for an opinion about the truth of a proposition” (Cruttenden 1997) or require affirmation or negation (Geluykens 1986). Additionally, zero-particle questions may be accompanied by paralinguistic features such as raised eyebrow, voice modification, and sometimes, body gestures particularly in languages that mark this question type solely intonationally (see Awobulu 1978).

The data analyzed in the present study show that the Yoruba yes/no question without particle (24) is characterized by an overall (tone by tone) upshift of the whole utterance on an extra-high pitch than for the corresponding statement (19) (see Fig.4.6; compare also Figs. 4.1 and 4.7). Unlike in Hausa, where downdrift is

Figure 4.8: Duration chart for STM (19). Except for Informant OPEM, the penultimate syllable in the utterance seems reliably longest for the informants.

Figure 4.9: Duration chart for Z-PARQ (24). The last syllable is the longest in duration for all the informants.
suspended\textsuperscript{25} (Miller and Tench 1980:54), downdrift is not suspended in this question type in Yoruba; it is compromised internally with a narrower percentage fall and a lower mean absolute slope than for statement (see Table 3). It is uttered on a faster tempo thereby having a shorter duration than the statement. The pitch range, as depicted in Fig. 4.6, is narrower between the two utterances at the onset than utterance-finally. The wideness, utterance finally, indicates a raising of the final syllable that was reported for Hausa (Miller and Tench 1980:52; Lindau 1986:761); for Seoul Korean (Bae 1998), and for French (Di Cristo 1998:203). The final syllable is longer in

\begin{figure}[h]
\centering
\includegraphics[scale=0.5]{fig410.png}
\caption{Intonation graph for question-word initially in Q-word question (25).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[scale=0.5]{fig411.png}
\caption{Intonation graph for question-word finally in Q-word question (26). Final syllables are raised. See Table 4.2 for actual difference in Fo values between the last two H syllables.}
\end{figure}

\textsuperscript{25} Suspension is technically used to mean “disobedience to” or “non-conformity to” the phenomenon under discussion. As reported by Miller and Tench (1985:54) for example, Hausa zero-particle questions do not obey the rule of downdrift.
duration when compared to the penultimate syllable, which in a statement appears longer (cf. Figs. 4.8 and 4.9).

Except in pedagogic textbooks (e.g. O’Connor and Arnold, 1973) where high-rise is consistently maintained as the intonation of English yes/no question (see also Hirst 1998:64), British linguists’ opinions differ on the nature of yes/no question (see Kingdon 1958; Halliday 1968; Crystal 1969; Geluykens 1986; Cruttenden 1997). In most tone languages on the other hand, overall higher pitch of the whole utterance seems to be the major factor for identifying yes/no question without particle. Other authors who have confirmed a similar occurrence in Hausa are Newman (2000) and Jaggar (2001). For Chinese, DeFrancis (1963: xxx) reports that all questions are uttered on a higher upshifted register, regardless of the kind of answer expected. The Chinese case is almost similar to Yoruba questions, except that though questions are generally higher in pitch than statements, range varies by question type.

Figure 4.12: Intonation graph in two-syllable-question-word initially in Q-word question (27).

Figure 4.13: F0 contours for Informant JIDM in sentences:
Ara won ni Opara. (All-M Statement = Lowest contour);
Ara won ni Opara? (All-M Zero-particle question = Topmost contour [similar shape with the lowest contour]);
Ara won ni Opara bi? (Contour ends with “bi?” to show the contour for All-M with a question word H “bi?” utterance-finally;
Se ara won ni Opara? (Contour starting with “Se”, i.e. All-M with H (se) as question word; and, Sikira lo mowo. (All-H tone statement contour with added texts).

In this figure, the F0 contour of All-M zero-particle question is higher than the All-H statement. The H particles in the questions are also seen to be considerably higher than the F0 contour of the All-H statement.

### 4.2.1.2. Question-word question

In this section, the nature of question-word questions is explored. Two types studied include question-word utterance initially (sentences 25 and 27), and question-word utterance finally (sentence 26). All the three, as used in this study, are pragmatically similar in terms of the type of responses they can elicit. While the question words in (25) and (26) are one-syllable H tone words, (27) is a two-syllable question word (Abi $\rightarrow$ LH). That the question words in (25) and (26) are H tone words does not suggest that Yoruba lacks one-syllable question-word of low or mid tones; the H tone types seem more common hence the choice in this study.

(a) Schematic representation of F0 contours in one statement and two questions (All-H tone sequence).

(b) Schematic representation of tonal contour in one question (see tonal sequence below).

Figure 4.14: F0 contour for Informant TOLF: (a) Sikira lo mowo. (Lowest contour) Se Sikira lo mowo? (Middle contour) Sikira lo mowo bi? (Topmost contour) (b) L H HHH H H H (Abi Sikira lo mowo?)

Different contours generated by All-H questions in this figure are interesting features of superimposed intonation. All the questions are raised utterance finally.
The analysis shows that the question-word question in Yoruba behaves more like the Hausa yes/no question, where a higher pitch of register is accompanied by a terminal rise (see Figures 4.11 and 4.14). To some extent, downdrift/declination is compromised, but the presence of, and the environment in which question words occur in the utterances make downdrift potentially suspendable. While the initial H tone, i.e. the question word in Figure 4.10 is “abnormally” high for onset Fo compared to statement, a comparison of the last two syllables in Figures 4.3 (all H statement) and 4.11 (question) shows that downdrift/declination is effectively blocked utterance finally in the question, i.e. in Figure 4.11, so that the intonation contour featured here is similar to the type described for Hausa. It is to be noticed also that though Connell and Ladd reported a slight rise utterance finally for Yoruba all H tone statements, the similar rise for three of the four informants (JIDM, COMF, and TOLF) in my data is absolutely negligible (see Table 2). For the all-H statement (Fig. 3), the difference in Fo between the penultimate and the final syllables are 1.04Hz, -0.37Hz, 1.56Hz, and 0.73Hz respectively.

![Figure 4.15: Intonation graph for Wh-clause question (28).](image)

![Figure 4.16: Intonation graph for Wh-clause question (29). There is lowering of di (H) at clause boundary for informants COMF and JIDM. This seems to be intonationally significant; particularly when the levels of OPEM’s and TOLF’s raising of the same tone are considered.](image)
For the question-word question ((26), Fig.4.11), the informants recorded a significant difference of 2.3Hz, 11.5Hz, 12.4Hz, and 16.3Hz respectively between the last two H tones. The higher Fo values recorded for the differences between the last two syllables of the question-word question by the informants is nothing but the effects of a superimposed intonation. The rise in pitch is due more to the fact that the final syllable is a question word. The same goes for the “abnormal” high Fo utterance initially in Figure 4.10 for instance. There is no indication that the Fo value for the initial H tone would be that high in a statement. The theoretical implication of all these points is that, question-words, particularly H tones, in the initial position are likely to trigger higher Fo onsets, and consequently raise the register of the whole utterance. More research is definitely required in this respect.

Though Figures 4.13 and 4.14 do not form part of the data sets in this study, they nevertheless provide copious illustration of how downdrift could be blocked in Yoruba question-word questions. Fig. 4.13 also shows an instance of overall upshift in register for the all M tone zero-particle question. The pitch height for the all H tone statement (Sikira lo mowo (21)) in the figure similarly helps to establish that questions are higher in pitch than statements. Figure 4.14 provides further evidence that Yoruba question-word question of yes/no type is raised in pitch utterance finally. Further, the claim by Connell and Ladd (1990), and Laniran (1992.) that the sequence LH does not trigger downdrift is confirmed in Fig. 4.14(b).

<table>
<thead>
<tr>
<th>Informants</th>
<th>Last two syllables</th>
<th>Tone Type</th>
<th>All-H (21)</th>
<th>Diff. in FO</th>
<th>QPARQ2(26)</th>
<th>Diff. in FO</th>
</tr>
</thead>
<tbody>
<tr>
<td>JIDM</td>
<td>mo wo</td>
<td>H H</td>
<td>173.73</td>
<td>1.04</td>
<td>208.8</td>
<td>211.1</td>
</tr>
<tr>
<td>COMF</td>
<td></td>
<td></td>
<td>253.41</td>
<td>-0.37</td>
<td>199.3</td>
<td>210.8</td>
</tr>
<tr>
<td>OPEM</td>
<td></td>
<td></td>
<td>201.99</td>
<td>1.56</td>
<td>163.5</td>
<td>175.9</td>
</tr>
<tr>
<td>TOLF</td>
<td></td>
<td></td>
<td>309.53</td>
<td>0.73</td>
<td>282.7</td>
<td>299</td>
</tr>
</tbody>
</table>

Table 4.2: Comparison between the Fo values of the last two H syllables in All-H statement (21) and Q-word question (26).

As seen in Table 3, the two-syllable question-word question (“Abi”— LH [QPARQ3]) is highest of the three question-word utterances in overall pitch for the extracted “omo naa ti de”, yet, internally, downdrift is compromised. So the expectation that the question should be higher in pitch is true for all the informants. As mentioned earlier, however, there is no clear reason for a LH-initiated syllable utterance to be higher than a H-initiated one because, ordinarily, one would expect the initial L tone to exert a pulling-down effect on the whole utterance just as H tone might be expected to exert a pulling-up effect. Within the extracted (i.e. “omo naa ti de”), the compromised downdrift is narrower in slope for informant TOLF, but wider for the other three informants (see Figure 4.12).
Table 4.3: Details of intonational variations between the utterances.

<table>
<thead>
<tr>
<th>Utterance Type</th>
<th>STM</th>
<th>Z-PARQ</th>
<th>Q-PARQ1</th>
<th>Q-PARQ2</th>
<th>Q-PARQ3</th>
<th>WH-CLQ1</th>
<th>WH-CLQ2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Pitch Height (Overall)</td>
<td>193.8</td>
<td>258.6</td>
<td>226.1</td>
<td>221.3</td>
<td>222.6</td>
<td>200.5</td>
<td>211.1</td>
</tr>
<tr>
<td>Average Pitch Height (Extracted)</td>
<td>193.8</td>
<td>258.6</td>
<td>216.1</td>
<td>222.6</td>
<td>230.6</td>
<td>193.3</td>
<td>202.9</td>
</tr>
<tr>
<td>Mean Absolute Slope (Semitone/s)</td>
<td>37.97</td>
<td>31.75</td>
<td>29.468</td>
<td>30.94</td>
<td>27.78</td>
<td>31.375</td>
<td>28.78</td>
</tr>
<tr>
<td>Percentage Narrowing</td>
<td>14.14</td>
<td>12.42</td>
<td>10.68</td>
<td>.</td>
<td>11.67</td>
<td>12.08</td>
<td>7.32</td>
</tr>
</tbody>
</table>

Figure 4.17: Bar-chart showing average pitch heights of the utterances. While the STM records the lowest Fo (193.8Hz), the Fo values for the Wh-clause questions come next. This is followed by the Q-word question values, which also rank about the same level on the chart. The zero-particle (yes/no) question maintains the highest Fo averagely. This shows that the yes/no questions without particle require the highest pitch of voice to utter them; otherwise the listener may fail to process them as questions. In other words, the extra pitch of voice is a demonstration of superimposed intonation.
Average Pitch Height for Extracted "omo naa ti de"

This figure was drawn for a planned comparison of the pitch heights of “omo naa ti de” in all the test items. The hypothesis being tested here is, if Yoruba intonation is a mere occurrence of lexical items in sequence, then the “Extracted” must have similar values all through. The hypothesis is refuted because the results show different Fo values for different utterances. Interestingly too, the values for similar utterance-types (Q-word questions QPARQ1, QPARQ2, and QPARQ3; and Wh-clause questions WH-CL1 and WH-CL2) are numerically close as in Figure 4.17. This confirms a superimposed intonation on Yoruba lexical tones.

4.2.2 Wh-Clause Question

Table 3 and Figures 4.15 and 4.16 provide the bases for discussing Wh-clause questions. Results show that Wh-clause questions in Yoruba are also said overall on a pitch slightly higher than, but very close to statements. Compromised downdrift is evident in the Wh-clause up to about the middle of the utterance in Fig. 4.15. Towards the end downdrift is suspended: contrary to the expectation that after the first high tone in an utterance, subsequent highs will lower progressively, it is observed in

Figure 4.19: Bar chart showing the percentage narrowing between the AP and EP of the “Extracted” in the utterances. The narrowing could not be determined otherwise because of tonal variation as well as length differences between the utterances.
Figure 4.15 that the H of na has a higher Fo than the H of the preceding pe. This may be due to the effects of clausal relationship intonation. Similarly, the expectation that L will lower following H and M is not true at the clausal boundary in this Figure. Therefore what happens is a clause boundary effect on tone rather than ordinary tone effects. The same is true for the clause boundary in Figure 4.16, where the Fo value of the Wh-clause is depressed at di(H), and raised suddenly at following ti(H). Rather than assume normal effects of tonal interaction in this environment, I argue that what happens here is more intonational. The situation becomes more intonational than tonal by the variation at the clausal boundary in Fig. 4.16 between the informants for i and di (idi → LH). While informants JIDM and COMF lower the H after L, OPEM and TOLF raise theirs but all of them raise H at the following ti (H) — the onset of the main clause. This raising of core clause Fo gives a pattern similar to 4.15. This gives a theoretical impression that, certainly,
Yoruba has specific intonation pattern peculiar to clause boundaries, but clausal boundary intonation is beyond the scope of this study. I expect to present an in-depth discussion on this when necessary works on it are completed.

In classificatory terms, Wh-clause questions are quite close in pitch values to statement (see Table 3 and Figure 4.18). The Fo values for the statement and the two Wh-clause questions are numerically close (i.e. 193.8Hz for statement; 193.3Hz and 202.9Hz respectively for Wh-clause questions). The closeness too of their heights on the bar chart indicate that they are better grouped together. If so, then Wh-clause question in Yoruba can be compared with Wh-question in English which has been reported to be said often on a falling tune. (e.g. Geluykens 1986:53; Ladd 1996:24, reporting Scherer et al. 1984).

The pitch values of the Yoruba Wh-clause are somewhat like those of the statement; between statement and question-word question actually (see Table 3; also Figures 4.17 and 4.18). I am assuming in addition that if the average pitch is this low when the Wh-clause starts with a high tone, the chances are high that when M or L tone is substituted in place of the initial H tone of this question type, the overall pitch may drop lower. That may then make the WH-clause question to look like a statement. At this point, the status of Wh-clause question is inconclusive but every indication shows that in Yoruba, it would likely be said on a falling tune.

Figure 4.19 shows the rate at which downdrift narrows in the utterances. Clearly, statements have wider range of declination than questions. Theoretically, downdrift is a phenomenon more common in statements than in questions. I do not claim to have completely mapped out the basis for the narrowing, but these findings are pointers to the fact that the intonation of statements is different from the intonation of questions in Yoruba. Furthermore, as many as 66% of the questions exhibit longer duration for the utterance final syllables (see Figures 4.20 through 4.22); statements lack such consistency so that it is difficult to make any such categorical statement.

5. Conclusion

Analysis reported for the elicited data in this study shows clearly that Yoruba does indeed permit a superimposed intonation on its lexical tones. As a matter of fact
the intonation of statement is distinct from the intonation of question in the following ways:

- Statement intonation is characterized by a downward progressive lowering (downdrift) of the pitch of voice in the course of the utterance, i.e. falling tune.
- Questioning intonation, on the other hand, is characterized by features which include: (1) general upward range-shift of the whole utterance register; (2) compromised downdrift, which narrows relative to question type; (3) final syllable lengthening; and (4) local rule suspension in some question types, i.e. some question types can effectively block the rule of downdrift. An interesting example is the concave and ladle-like pitch contours generated for the all high tone questions “Se Sikira lo mowo?” and “Sikira lo mowo bi?” respectively in Figure 4.14a.

Furthermore, some questions, e.g. yes/no question without particle and question-word question raise the last syllables (Figure 4.14 shows this more conspicuously). For a greater percentage of the questions, the final syllable lengths considerably in duration. At the superficial level, one may want to assume that the lengthening is occasioned by the final raising. The fact, however, that the zero-particle question, which shows a remarkably overall higher pitch than its statement counterpart, was uttered on a shorter final-syllable duration makes it very difficult to make such assumption.

Based on my findings, I classify Yoruba intonation, in ascending order, into four pitch levels— Level 1 (the lowest) for Statements; Level 2 for Wh- questions; Level 3 for Question-word Questions (or yes/no questions with particle); Level 4 for Yes/No Questions (without particle). These phenomena constitute intonation in Yoruba.

6. General Conclusion/Suggestion

- Since the distinctiveness of tones in tone languages do not make tones intonational in any fixed manner, I reiterate again that there is no such language as an intonation language. Any language can be either a stress language or a tone language; and the two systems permit a superimposed intonation, even though tone languages do not seem to require extensive use of intonation as does stress languages like English for instance.
- Conventionally, downdrift/downstep has remained the sole phenomenon for describing tone language intonation, but I argue that the phenomenon works only for statement intonation in the form of declination. If it can be effectively blocked by some questions, then some additional phenomenon or phenomena (or framework(s)) need to be proposed for questioning intonation in tone languages.
- Pitch and duration have proved more useful in the present analysis. The claim that intensity correlates very well with fundamental frequency (e.g. Zee 1978) is only partially supported by the findings in this study. There is no consistency in the way it works. In some environments, high pitch attracts an equally high intensity. But in some others, lower Fo take higher intensity. The reason is not clear yet, but perhaps, intensity bears some special relationship to tone types. This, no doubt, requires further exploration.
7. Suggestions for Further Studies

The model which has been developed for use in this study has proven instrumental to the discoveries made. In the model, it has been possible to map out global intonation graphs from the underlying local rules of tone without necessarily having to place targets within two grid lines as did previous researchers, who have worked on two-toneme languages. The generated “curves” are self-explanatory to a large extent. One other advantage of the model is its ability to separate the rules of tone from those of intonation. In other words, the model enables a clear distinction between tone (the underlying sequences of linguistically significant local targets H, M, and L) and intonation (the generated global Fo pattern). It is therefore expected to enhance further studies of this nature. Such areas needing further research in Yoruba include: imperatives, attitudinal intonation, intonation of vocatives, discourse intonation, focus, and the intonation of clause relations. And maybe at some point, the tonal contour problems resulting from tone interaction can be effectively handled to analyse tone language intonation using the head-types, nucleus-position, tail etc as has conventionally been used to describe English for instance; or at least in a way compatible with such a type of description.

REFERENCES


26 I am assumption that if the model proved useful thus far, it might worth trying it on these phenomena as well.


Fraser, Bruce. 1977. “The role of intonation in signifying speaker intent” In Fasold, R.W. & Shuy, R. W. (ed.) *Studies in Language variation: semantics, syntax, phonology, pragmatics, social situations, ethnographic approaches*; Georgetown University, USA.


