An Acoustic Study of Vocal Emotions in Yoruba

Kolawole Adeniyi

Department of Linguistics and African Languages, Obafemi Awolowo University, Ile-Ife, Nigeria
kola.adeniyi@oauife.edu.ng

Abstract

The vocal expression of different emotional states is an essential part of communication. However, vocal emotion in African languages has not attracted much scholarly works. This study investigates the roles of pitch, intensity, and speech rate in the vocal expression of angry, happy, sad, fear, disgust and neutral emotions in Yoruba, a West Benue-Congo language spoken predominantly in Southwestern Nigeria. Data were gathered from 12 professional and church actors using utterances that can be expressed in the six different emotional states. The results of the acoustic analysis show that the expression of emotions in Yoruba employs a blend of cues from pitch level, intensity and speech rate. Although the presence of three tones plus a downstep in Yoruba results in a congestion of the tone space thereby placing a limitation on the effects of pitch change in the discrimination of vocal emotions, notable differences in pitch change still exist across the emotional categories. It is concluded that the vocal expression of emotion deploys a complex interaction between pitch, speech rate and intensity.

Keywords: *vocal emotion, pitch, speech rate, intensity, actors*

Introduction

The primary function of emotion is to mobilize the organism to deal quickly with important interpersonal encounters (Ekman, 1999). However, the mechanisms (especially articulatory and cognitive) behind this speech activity are yet to be satisfactorily deciphered. The study of emotion is of interest to neurologists, psychologists, computer scientists, artists and linguists among others. Relating to the emotional states studied in this work, Ekman (1999, pp. 46) posits that happy emotion entails the attainment of a goal; sad, "failure to attain or maintain a goal"; angry, loss of a goal; fear implies expectation of failure; and disgust is response to 'something offensive to taste or smell'.

Since emotion is an essential part of human communication, a good understanding and deployment of emotion enhances communication (Ekman, 2003). It can be conveyed through speech, facial expression or body language. The vocal expression of emotion remains a vital part of speech (Jürgenset al., 2011; Preti, et al., 2016). In this age of extensive human-machine interface in all facets of human life, it is essential that emotion-modelling be a part of speech modelling to ensure that the output of machines will suit users' emotional states, and also to ensure that machines are able to decode emotional states in natural voice and process accordingly.

The understanding of cross-cultural processing of vocal emotions remains elusive as the exact roles of pitch, intensity and speech rate are yet to be precisely articulated. For instance, some studies claim that emotion recognition is largely universal (Ekman, 2003; Gussenhoven, 2004; Wang & Lee, 2015), while oth-

ers note variations in emotion cues across languages. Urua and Ekpenyong (2020) note that there is no agreement on which speech features are sufficient to distinguish emotions. This reinforces the idea that since Acoustic cues of emotion vary across languages (Breitenstein et al., 2001; Wang & Lee, 2015), African Tone Languages (ATL) need to be studied in order to key fully into the full benefits of the speech modelling component of human-machine interface.

Literature Review

A number of studies have appeared in the literature with focus on the vocal expression of emotions. For instance, Jürgens et al. (2011) studied the influence of play-acting on vocal emotions by comparing "radio sequences of naturally occurring emotions to actors' portrayals" in German-speaking subjects, and reported variations between the voice qualities of the two stimuli. Specifically, they report that play-acted speech showed more variable pitch-contours, which indicates differences in the encoding of emotion between natural and play-acted speech. Jones et al. (2011, p.982) studied the emotion of anxiety in the speech of 63 Greek speakers "in a situation which induced mild anxiety" and report that anxiety showed a significant increase in mean fundamental frequency and in noise-to-harmonic ratio.

Preti et al. (2016) studied the emotional states of angry, happy, and neutral in two Italian actors. Their result showed that fundamental frequency, vocal intensity, and speech rate indicate individual perceptions of the three emotions. Also, Wang and Qian (2018, p.324) conducted production and perception experiments to determine whether lexical tone could influence the ability of pitch to project emotion in Mandarin-Chinese and concluded that "listeners may not always use pitch variation cues to identify emotions in speech".

In their study of fear, disgust, joy, sadness, hot anger and cold anger emotions, Probst and Braun (2019) recorded six professional German actors who produced each emotional state in three degrees of low, medium and extreme. The analysis was done manually on PRAAT. They report that speakers use modulations of their fundamental frequency to express emotion, while "[i]ncreasing degrees of emotionality are in fact implemented by increasing mean f0 for the portrayal of emotions" (Probst & Braun, 2019, p.68). Also, Yildirim et al. (2004) investigated the acoustic properties of speech associated with sad, angry, happy, and neutral emotions intentionally expressed in speech by a single actress. They measured "speech prosody, vowel articulation and spectral energy distribution" (p.2193) and report that "[r]esults show that happiness/anger and neutral/sadness share similar acoustic properties in this speaker" (p.2196). According to the authors, the parameters studied did not effectively distinguish between angry and happy in the study.

On African languages, Urua and Ekpenyong (2020) conducted a Principal Component Analysis study of the effects of duration, pitch, intensity and formants on seven emotion types, namely angry, fear,

joy, normal, pride, sad and surprise in Ibibio and reported that formant features were "most responsible and useful for discriminating variability profiles" in African Tone Languages. One area of contention with this work is the variability in the utterances used to test the different emotional states, which is capable of influencing acoustic cues. This is because the utterances were of different types (declarative, interrogative, etc.) and lengths. Also, Cahill (2017) reports that the expression of emotion is intonational in Kɔnni, a Ghanaian language. In their study of vocal expressions in Yoruba poetry, Akinbo et. al (2022) report that pitch and vibrato were reliable markers of distinction between *Èsà* and *Ìjálá*, two similar genres of Yoruba poetry. Akinbo et. al (2022, p.12) note that "studies on effective use of vocal expression find that pitch raising and increased loudness are the most reliable cues for high level of arousal, such as excitement, fear and anger."

Problem Statement

Studies on emotion have been confronted with some general challenges. One problem that studies on emotion are fraught with is access to natural data; different methods ranging from the use of questionnaires to play-acted utterances have been deployed. There is then the methodological problem of the need to draw a line between a genuine modelling of emotion and mere theatrics. Some other methods of data elicitation have been deployed to suit different languages, but in Yoruba, a three-tone language with a more relative congestion of the tone space, there is need for methods that suitably capture the specific speech dynamics in order to fully understand the acoustic cues of vocal emotions and the interactions between such cues.

Also, most researches on vocal emotion have been largely statistical, leaving a gap in the details of the linguistic aspect, and this calls for extensive linguistic study of vocal emotions in African tone languages. These challenges are taken up in this study. This present study is, therefore, an investigation of the roles of pitch, intensity and speech rate in the vocal expression of angry, happy, sad, fear, disgust and neutral emotions in Yoruba, a Nigerian tone language. In addition to the roles of vocal emotions in communication, the enunciation of the mechanisms that Yoruba uses to express vocal emotions is significant to pedagogy since it adds to understanding of what needs to be taught, and reveals crucial aspects of teacher-pupil relationships. The study therefore sets out to examine the roles of pitch, intensity and speech rate, as well as the interaction between these emotional cues, in the vocal expression of emotion. This is to provide a linguistic platform for emotion modelling in Yoruba.

Background on Yoruba

Yoruba is a West Benue-Congo language spoken in Nigeria's south-western states of Lagos, Ogun, Oyo, Osun, Ekiti, and Ondo, as well as parts of North-central states of Kwara and Kogi. As of 2015, Yoruba speakers in Nigeria numbered around 39,500,000 people (Simons & Fennig, 2018). At the rate at which Nigeria's population is growing, speakers of Yoruba are now well over 40, 000, 000. Yoruba is also spoken as

a native language in the Republic of Benin, Togo, Cote d'Ivoire, Ghana, Sierra Leone, Brazil and Cuba (Adetugbo, 1982; Fabunmi, 2013; Oyetade, 2011).

Yoruba is a tone language, which operates three lexical tones (High < ´ >, Low < ` >, and Mid, which is usually left unmarked), two contour tones (rising tone < ˇ > and falling tone < ˆ >), downstep and downdrift (Bamgbose, 1967; Connell & Ladd, 1990; Adeniyi, 2020). Pitch, which is one of the foci of this study, is the perceptual correlate of tone. It has been extensively studied in Yoruba, albeit with attention paid to either tone (Adeniyi, 2020; Connell & Ladd, 1990), or intonation (Laniran, 1992; Laniran & Clements, 2003). In this study, pitch is viewed in relation to its role in the vocal expression of emotion in Yoruba.

Methodology

This study employed a mixed-method approach which allowed for an experimental collection of data considered reliable for acoustic analysis (See Abbuhl et al., 2014, pp. 124-127 for details on mixed-method approaches). The method is detailed in the following sub-sections.

The data and participants

Data were sourced through production experiments requiring 12 professional and church actors to produce specific utterances depicting different emotional states. It was ensured that the participants were all people who conducted their everyday lives in Yoruba, to exclude interference from other languages. The experimental participants were constituted into two focus groups and each member of the groups produced a set of purposively structured utterances in the required emotional frames. The consent of the production experiment participants was obtained before recording and those that declined were not recorded but they participated in the perceptual scrutiny aspect of the recordings.

Declarative utterances that could be said with different emotions depending on circumstances were used. This requirement to use utterances that can be said in the different emotional states is similar to Yildirim et al. (2004). Many of the utterances were drawn from societal matters that divide opinions within the contemporary Yoruba society. Each of the utterances (*i-vii*) was said three times by each of the production experiment participants.

Test materials

(i) **Bíléyọrí   ti   lóyún**

Bileyori   is   pregnant

"Bileyori is pregnant"

(ii) **Ẹran ajá   ló   ńjẹ**

Meat dog   that.he   ASP.eat

"it is pock that s/he is eating"

(iii) **Bíléyọrí   ńjó**

Bileyori   ASP.dance

"Bileyori is dancing/burning"

(iv) **Ọmọ kan   ṣoṣo   tí   Bíléyọrí   bí**

Child one   only   that   Bileyori   give.birth.to

"Bileyori's only child"

(v) **Ilé   Lámíkan   ni   Bíléyọrí   wà**

House Lamikan   that   Bileyori   be

"Bileyọri is in Lamikan's house"

(vi) **Bíléyọrí   gbọ́mọ   lánà-án**

Bileyori   carry.child   yesterday

"Bileyori carried baby yesterday"

(vii) **Ọjọ́   òní   yàtọ̀**

Day   today   different

"today is different"

Perceptual scrutiny of data

The presence of the focus group members during the recording sessions served as control and allowed for the scrutiny of each utterance produced. The utterances in *i-vii* were taken one at a time and each participant produces them in the six emotional states. Utterances produced were then subjected to perceptual scrutiny whereby participants were required to judge whether all the emotional states were accurately and clearly represented. It was only the utterances that every member adjudged as depicting the required emotional state that were passed for acoustic analysis. An expression of doubt over the accuracy of one emotional state (in a set of six tested) rendered the entire six unusable for the affected utterance. This was done for all the utterances and all the participants were allowed to repeat their production as often as they wanted.

Recording

Main recording sessions were preceded by series of practices aimed at making the atmosphere as natural as possible. It was not possible to gather the participants into a sound-proofed room, so recordings were done in quiet rooms during the periods of general quietness in the environment. There was an average of six peo-

ple present at each recording session. Recordings were done with Zoom H1N digital voice recorder at 44 kHz sampling rate at 16 bits resolution.

Each focus group was allowed to determine how to be prompted to produce the utterances in the required emotional states. For instance, while one group created mental scenarios depicting each emotional state, the second group used play-acted dialogues as prompt for each of the emotional states. Participants were also allowed to practice for as long as was required before recordings could be done.

Analysis procedure

The acoustic analysis involved the measurement of pitch, intensity and speech rate. Measurements of each utterance were taken globally (spanning entire utterances), since we were able to use the same segmental materials for all the emotional states tested. The use of the same segmental materials for all the emotion types eliminated other factors that could interfere with the results. Data were processed utterance-by-utterance. Under each utterance, the six emotions were collated.

Where a participant failed to produced one emotion for an utterance, that participant's tokens were disqualified for analysis of such utterance. Pitch, intensity and speech rate were measured in each utterance and scores were tabulated in an excel spreadsheet. The recorded scores were then compared across the utterances to see the individual as well as combined effects of the variables on vocal emotions in Yoruba.

Findings

A key component of the result is the fact that all the participants in the experiment showed strikingly similar patterns for all the emotional states tested. This similarity in patterning is consistent across the three variables of pitch, intensity and speech rate in the tokens of all the participants. This allows for a qualitative examination of a sample utterance. Table 1 contains the average values of three participants, two female and one male for ***Bíléyọrí ńjó*** *"Bileyori is dancing/burning."* It will be seen that neutral emotion recorded the lowest value for pitch and intensity, while angry had the highest value for pitch and intensity, and was also the fastest. Pitch is however the most prominent indication of angry emotion, where it is 11Hz higher than the nearest emotions. Happy and fear had similar pitch and intensity values; with the only difference between them being that fear is slightly faster. Sad and disgust had similar intensity, but the pitch of sad was13.2Hz higher than disgust whereas disgust is faster than sad. A more detailed discussion of the three factors is presented below.

Table 1: Average values for *Bíléyọrí ńjó* "Bileyori is dancing/burning"

Bíléyọrí ńjó "Bileyori is dancing/burning"	Pitch (Hz)	Intensity (dB)	Time (sec)
Angry	250.8667	79.46667	1.033333
Happy	239.8667	78.56667	1.566667
Sad	192.6333	75.1	1.533333
Fear	239.9667	78.03333	1.2
Disgust	179.4333	74.33333	1.233333
Neutral	170.3	72.33333	1.066667

Pitch as a cue for emotion discrimination in Yoruba

More specifically, pitch appears to be a clear indicator of vocal emotion in Yoruba. According to Adeniyi (2015), a pitch difference of 10Hz is usually perceptible to the ear in a three-tone language. Going by this, it can be seen that any form of vocal emotion entails a raise in pitch since neutral has the lowest pitch; emotion discrimination is then dependent on the degree to which the pitch is raised. Except for happy and fear that recorded similar values, all the emotional states tested have significant pitch differences.[1] This results in a five-way discrimination (happy/fear ~ angry ~ sad ~ disgust ~ neutral). The blue line in Figure 1 provides a visual presentation of the result.
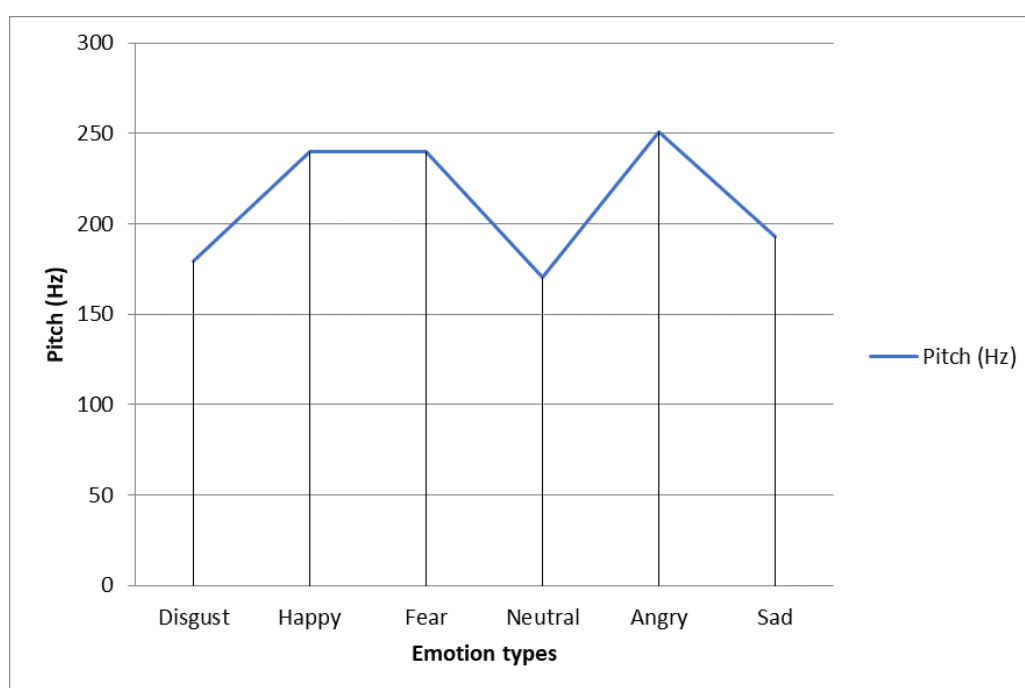

Fig. 1: Pitch as a cue of emotional state in Yoruba

---

[1] At 9.1Hz, the different between disgust and neutral is also slightly lower that the expected 10Hz.

Intensity as a cue for emotion discrimination in Yoruba

Table 1 also shows that verbal expression of emotion entails the raising of intensity, since neutral emotion recorded the lowest value. 1.3dB is regarded as the variation for which words are perceived as being different if intensity alone were considered (Rogers, Healy, & Montgomery, 2006; Orr, Montgomery, Healy, & Dubno, 2010, p. 3105). Going by this, all the non-neutral emotional states recorded significant difference from neutral (Fig. 2). However, it is shown that intensity-based discrimination between non-neutral emotions depends on the degree of raising for each emotion. In that sense, angry, happy and fear recorded values that were not significantly different from one another at the highest level, while sad and disgust were similar at a relatively intermediate level between neutral at the lowest level and angry, happy and fear at the highest level. This presents a three-way partial discrimination (angry/happy/fear ~ sad/disgust ~ neutral).
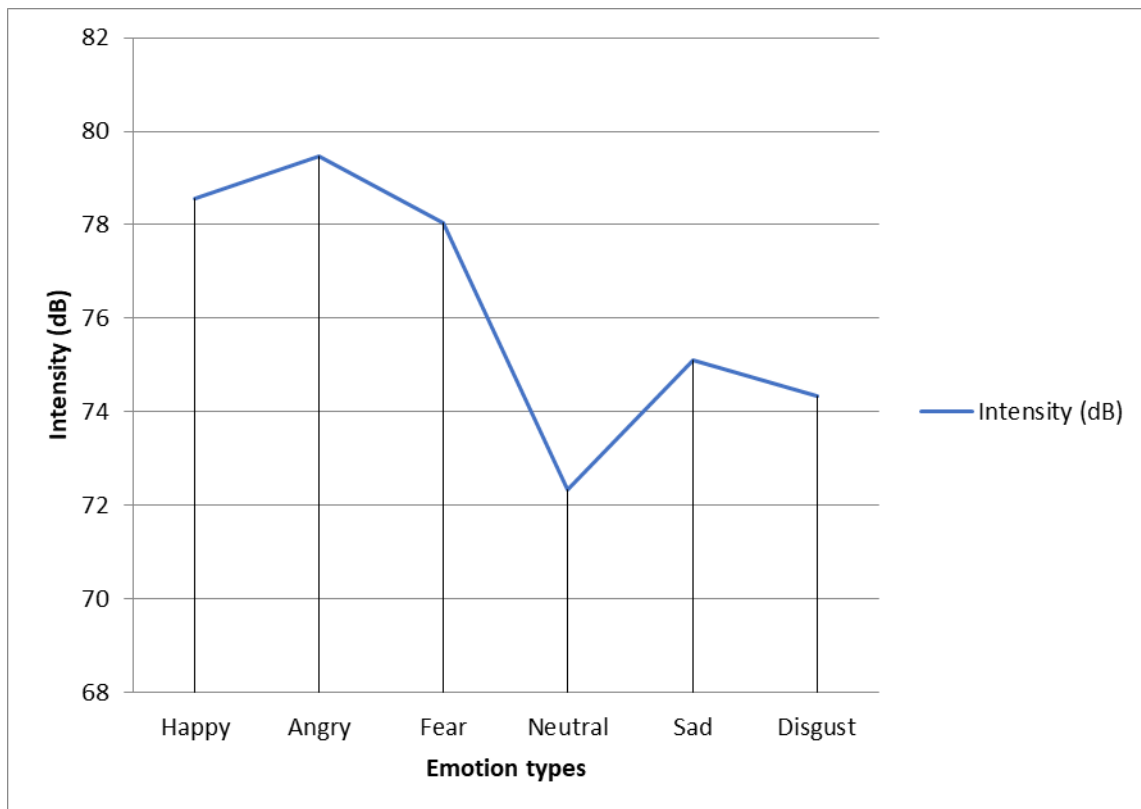

Fig. 2: Intensity as a cue of emotional state in Yoruba

Speech rate as a cue for emotion discrimination in Yoruba

From the result in Table 1, only angry was produced in a shorter time than neutral emotion. This is consistent with native speaker intuition that Yoruba speakers speak faster when angry. Happy and sad emotions were the slowest at about 1.5 seconds, which is also in consonance with native speaker intuition that people may not speak as fast when happy or when sad, but it must be noted that these two emotional states are not

distinguishable on the basis of speech rate. Fear and disgust were also not distinguishable on the basis of rate of speech as both recorded a speed of about 1.2 seconds.

Categorization of emotional states

It is apparent that none of pitch, intensity or rate of speech is independently sufficient in discriminating between the tested emotional states in Yoruba. An attempt is made in this section to see whether a clear distinction between the six emotional states can be gotten via the interaction between the three factors of pitch, intensity and speech rate. Table 2 contains comparative/superlative categorizations of the emotional categories in this wise. In determining the different categories, a difference of 10Hz and above was considered significant for pitch (Adeniyi, 2015), a difference of 1.3dB and above was considered significant for intensity (Orr et al., 2010; Rogers et al., 2006), while a timing difference of 0.2 seconds and above was considered significant for rate of speech.[2]

Table 2, which appeared highly simplified, shows that none of the six emotional states has the same value set with another. A closer examination reveals further that in spite of using only four dimensions for the categorization, each emotional state differs from the others in at least two dimensions. This suggests an evident distinction between the six emotional states on the basis of interaction between pitch, intensity and rate of speech.

Table 2: Properties of emotions compared to neutral in Yoruba

Bíléyọrí ńjó "Bileyori is dancing/burning"	Pitch (Hz)	Intensity (dB)	Time (sec)
Angry	highest	loudest	neutral
Happy	high	loudest	faster
Sad	higher	loud	faster
Fear	high	louder	fast
Disgust	neutral	loud	fast
Neutral	neutral	neutral	neutral

The result of this study confirms some of the tendencies reported in some previous studies reviewed, while it differs from others. For instance, the fact that neutral emotion has lower f0 than other categories is consistent with works reviewed for this research (see for instance, Probst & Braun, 2019; Yildirim et al., 2004 among others). But whereas Yildirim et al. (2004) report duration similarity between angry and happy

---

[2] The categorization is similar to Cahill (2017).

emotions, and f0 value similarities between sad and neutral, the report of this study suggest otherwise. This is because angry is 0.54 sec slower than happy, and sad recorded f0 that was 22 Hz higher than neutral. Also, Probst and Braun (2019) note that Joy, fear and anger "cause speakers to produce a higher mean f0, greater f0 variability and greater f0 range than *neutral* utterances do," which is consistent with the findings in this study (see also Yildirim et al., 2004). It is also worth noting that even with differences as low as 4Hz between emotions, Yildirim et al. (2004) report a statistically significant impact of emotion on mean fundamental frequency. Their results were 188Hz for neutral, 195 Hz for sad, 233Hz for angry, and 237 Hz for happy (Yildirim et al., 2004), but this study reports wider f0 differences across the emotion categories.

Discussion

The roles of pitch, intensity and speech rate in the vocal expression of six emotional states, namely angry, happy, sad, fear, disgust and neutral emotions in Yoruba were discussed in this article. The three properties of pitch, intensity and speech rate were shown to be cues for emotional states in Yoruba. However, none of the properties was sufficient to exclusively discriminate between the six emotional states. Rather it was reported that emotion discrimination is achieved through the interactions between the three properties tested. It is therefore concluded that although pitch may appear to be the single most crucial carrier of emotion in Yoruba, adequate emotional discrimination is achieved through an intricate interaction between the three phonetic properties tested. The implication of these results is that studies on emotion discrimination need to test as many emotional states as possible. Testing only a few emotional states may yield results falsifiable with studies with expanded scope. Also, the ability to use similar utterances to determine emotional cues is capable of yielding interesting results. It would be interesting to see the contribution of overall contour shapes to emotion discrimination in Yoruba. This is considered a direction for further studies on emotion discrimination in Yoruba.

Author contributions

Kolawole Adeniyi is the sole author of this paper. He is solely responsible for all aspects of the study and has authorized the published version of the manuscript.

Funding

None.

Data availability statement

Data is available on direct request to the author.

Conflicts of interest

There was no conflict of interest in this study.

References

Abbuhl, R., Susan G. & Alison M. (2014). Experimental research design. In R. J. Podesva, & D. Sharma (Eds.), *Research Methods in Linguistics* (pp.116–131). Cambridge University Press. https://doi.org/10.1017/CBO9781139013734

Adeniyi, K. (2015). *Downstep in three-tone Systems of West Benue-Congo Languages* [Unpublished Master's Thesis]. University of Ibadan.

Adeniyi, K. (2020). Lexicalization of tonal downstep in Yoruba. *Canadian Journal of Linguistics/ Revue Canadienne de Linguistique,* 65(4), 353–555. https://doi.org/10.1017/cnj.2020.22

Adetugbo, A. (1982). Towards a Yoruba dialectology. In A. Afolayan (Ed.), *Yoruba language and literature* (pp. 207–224). University of Ife Press.

Akinbo S. K., Samuel, O., Alaga, I. B., & Akingbade, O. (2022). An acoustic study of vocal expression in two genres of Yoruba oral poetry. *Front. Commun.,* 7,1-15. www.doi.org/10.3389/fcomm.2022.1029400

Bamgbose, Ayo. (1967). *A short Yoruba grammar*. Heinemann Publishers.

Breitenstein, C., Van Lancker D., & Daun, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition and Emotion,* 15(1), 57–79. https://doi.org/10.1080/02699930126095

Cahill, M. (2017). Kɔnni Intonation. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 53–88). Walter de Gruyter. https://doi.org/10.1515/9783110503524

Connell, B. (2017). Tone and intonation in Mambila. In L. J. Downing & A. Rialland (Eds.) *Intonation in African tone languages* (pp. 133–166). Walter de Gruyter. https://doi.org/10.1515/9783110503524

Connell, B. & Ladd, D. R. (1990). Aspects of pitch realisation in Yoruba. *Phonology*, 7(1), 1–29. https://doi.org/10.1017/S095267570000110X

Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. Power (Eds.) *Handbook of Cognition and Emotion* (pp. 45–60). Wiley Publishers. https://doi.org/10.1002/0470013494.ch3

Ekman, P. (2003). *Emotions revealed: recognizing faces and feelings to improve communication and emotional life.* Henry Holt.

Fabunmi, F. A. (2013). Negation in sixteen Yoruba dialects. *Open Journal of Modern Linguistics*. http://dx.doi.org/10.4236/ojml.2013.31001

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616983

Jones, M., Anagnostou, F. & Verhoeven, J. (2011). The vocal expression of emotion: an acoustic analysis of anxiety. *Proceedings of ICPhS XVII,* 982-985. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Jones/Jones.pdf

Jürgens, R., Hammerschmidt, K., & Fischer, J. (2011). Authentic and play-acted vocal emotion expressions reveal acoustic differences. *Frontiers in psychology, 2*. www.doi.org/10.3389/fpsyg.2011.00180

Laniran, Y. O. (1992). *Intonation in Tone Languages: The phonetic implementation of tones in Yoruba* [Unpublished Doctoral Dissertation]. Cornell University.

Laniran, Y. O. & Clements, G. N. (2003). Downstep and high raising: interacting factors in Yoruba tone Production. *Journal of Phonetics,* 31, 203–250. https://doi.org/10.1016/S0095-4470(02)00098-0

Orr, S. B. Montgomery, A. A. Healy, E. W., & Dubno, J. R. (2010). Effects of consonant-vowel intensity ratio on loudness of monosyllabic words. *Journal of the Acoustical Society of America, 128*(5), 3105–3113. https://doi.org/10.1121%2F1.3493426

Oyetade, S. O. (2011). The Yoruba language in diaspora: lessons from the English Language. *Yoruba: Journal of the Yoruba Studies Association of Nigeria, 6*(5), 1–24.

Preti, E., Suttora, C., & Richetin, J. (2016). Can you hear what I feel? A validated prosodic set of angry, happy, and neutral Italian pseudowords. *Behav Res,* 48, 259–271. https://doi.org/10.3758/s13428-015-0570-7

Probst, L. & Braun, A. (2019). The effects of emotional state on fundamental frequency. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences, Melbourne, Australia*, 67–71. https://www.assta.org/proceedings/ICPhS2019/papers/ICPhS_116.pdf

Rogers, C. F., Healy, E. W., & Montgomery, A. A. (2006). Sensitivity to isolated and concurrent intensity and fundamental frequency increments by cochlear implant users under natural listening conditions. *Journal of the Acoustical Society of America,* 119, 2276–2287. https://doi.org/10.1121/1.2167150

Simons, G. F. & Fennig, C. D. (Eds). (2018). *Ethnologue: Languages of the World* (21st Ed.). SIL International. http://www.ethnologue.com.

Urua, E. A. & Ekpeyong, M. (2020). Principal component visualization of acoustic-emotion profiles in Ibibio. In E. A. Urua, F. Egbokhare, O. Adesola & H. Adeniyi (Eds.), *African Languages in Time and Space: A Festschrift in Honour of Professor Akinbiyi Akinlabi* (pp. 338 -351). Zenith BookHouse.

Wang, T. & Lee, Y-C. (2015). Does restriction of pitch variation affect the perception of vocal emotions in Mandarin Chinese? *The Journal of the Acoustical Society of America, 137*. www.doi.org/10.1121/1.4904916

Wang, T. & Qian, Y. (2018). Are pitch variation cues indispensable to distinguish vocal emotions? *Proceedings of the 9th International Conference on Speech Prosody,* 2018, 324–328. https://www.isca-archive.org/speechprosody_2018/wang18c_speechprosody.pdf

Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A. Busso, C. Lee, D. Z. Z., &, Narayanan, S. (2004). An acoustic study of emotions expressed in speech. *Proceedings of InterSpeech*, 2193–2196. https://www.isca-archive.org/interspeech_2004/yildirim04_interspeech.html