



OPEN ACCESS

EDITED BY
Sandra Madureira,
PUCSP, Brazil

REVIEWED BY
Julien Meyer,
Centre National de la Recherche
Scientifique (CNRS), France
Alexandro Rodrigues Meireles,
Federal University of Espirito Santo,
Brazil

*CORRESPONDENCE
Samuel K. Akinbo
akinbosk@gmail.com

SPECIALTY SECTION
This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

RECEIVED 27 August 2022
ACCEPTED 28 November 2022
PUBLISHED 15 December 2022

CITATION
Akinbo SK, Samuel O, Alaga IB and
Akingbade O (2022) An acoustic study
of vocal expression in two genres of
Yoruba oral poetry.
Front. Commun. 7:1029400.
doi: 10.3389/fcomm.2022.1029400

COPYRIGHT
© 2022 Akinbo, Samuel, Alaga and
Akingbade. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

An acoustic study of vocal expression in two genres of Yoruba oral poetry

Samuel K. Akinbo^{1*}, Olanrewaju Samuel², Iyabode B. Alaga² and Olawale Akingbade²

¹Department of Linguistics, University of British Columbia, Vancouver, BC, Canada, ²Department of Linguistics and African Languages, University of Ibadan, Ibadan, Nigeria

This pilot study proposes an acoustic study of the vocal expressions in *ljalá* and *Ḃsà*, two genres of Yorùbá oral poetry. For this study, we conducted an experiment, involving the vocalization of an original poem in speech mode, *ljalá* and *Ḃsà*. The vocalizations were recorded and analyzed acoustically. The results of the study show that cepstral peak prominence (CPP), Hammarberg index and Energy of voiced sound below 500 Hz distinguish comparisons of *Ḃsà*, *ljalá* and speech but are not as reliable as F0 height and vibrato. By comparing the pitch trajectories of the speech tones and poetic tunes, we show that poetry determines tone-to-tune mapping but can accommodate language when it is feasible. The results of our investigation are not only in line with the previous impressionistic observations about vocal expression in Yorùbá oral poetry but contribute with new findings. Notably, our investigation supports vocal tremor as the historical origin of vibrato in *ljalá*. As a result of this, we strongly recommend the instruments of phonetic science for the study of vocal expression in African oral poetry.

KEYWORDS

oral poetry, tone, vibrato, vocal effort, vocal expression, phonetics

1. Introduction

One of the major challenges with the study of African oral poetry is lack of instrumental analysis, even when the analysis of vocal expression is based on terms with phonetic correlates (Babalola, 1963; Ajuwon, 1977; Olabimtan, 1977; Olatúnjí, 1979). In this preliminary study, we address this issue by proposing an acoustic analysis of vocal expression in two genres of Yorùbá oral poetry.

Studies show that vocal expressions, such as pitch raising and increased loudness, communicate emotions (Scherer, 1985; Juslin and Laukka, 2003). For instance, high pitch and increased loudness are associated with high level arousal such as anger, happiness and excitement (Banse and Scherer, 1996; Juslin and Laukka, 2003; Johnstone et al., 2005; Goudbeek and Scherer, 2010; Lindquist et al., 2016; Scherer, 2021). The use of vocal expression is not limited to affective communication. For example, in many African oral traditions, various genres of verbal art are distinguished based on the vocal expressions that are associated with them (e.g., Uzochukwu 1981 on Igbo elegiac poetry of Nigeria; Boadi 1989 on Akan praise

poem of Ghana; Brown 1995 on indigenous South African oral poetry Ọlátúnjí 1979 on Yorùbá oral poetry). The acoustic correlates of vocal expressions in affective speech have been extensively investigated (see the reviews in Scherer, 1986; Juslin and Laukka, 2003; Kamiloğlu et al., 2020), but the vocal expression in African oral traditions have mostly been analyzed without the instruments of phonetic science.

The present paper describes an acoustic study of vocal expressions in two genres of Yorùbá oral poetry. The study is based on Ịjálá and Ẹsà, two genres of Yorùbá oral poetry. We will argue that the instrument of acoustic phonetics can offer valuable insights on the verbal aesthetics of African oral poetry.

Before turning to the details of this study, we present the basic sound inventory of Yorùbá in Section 2. The discussion in Section 3 focuses on the features of Ịjálá and Ẹsà in the context of Yorùbá oral poetry. To understand the phonetic correlates of vocal expression in Ịjálá and Ẹsà, we conducted a production experiment. The details of the experimental procedure are presented in Section 4. The results of the experiment are presented in Section 5. The discussion and conclusion are presented in Section 6.

2. Background on Yorùbá sound inventory

Yorùbá is a Volta-Niger language spoken in West Africa and prominently South-Western Nigeria (Blench, 2019). This section presents a description of the relevant sound patterns in Standard Yorùbá, which is the basis of this work.

2.1. Tone in speech

Yorùbá is a tone language, which means pitch contrasts bring about lexical or grammatical distinctions in meaning (Yip, 2002; Hyman, 2018). The language has three contrastive tones, namely H(igh), M(id) and L(ow) (Akinlabi, 1985; Pulleyblank, 1986).

- (1) Yorùbá: Tonal minimal set
- | | | |
|---|----|------------|
| H | bá | 'to meet' |
| M | ba | 'to braid' |
| L | bà | 'to land' |

As shown in (1), H tone is marked with an acute accent and L tone with a grave accent. However, M tone is unmarked (Bamgboṣe, 1965; Awóbùlúyì, 1978). Throughout this work, this tone-marking convention in Yorùbá orthography is adhered to.

2.2. Vowels

The vowel inventory of Yorùbá contains seven oral vowels and three nasal vowels, which are presented in (2) (Bamgboṣe, 1966; Awóbùlúyì, 1978; Pulleyblank, 1988).

- (2) Yorùbá vowels (Pulleyblank, 2009, p. 868)

		Front	Central	Back
Oral	High	i [i]		u [u]
	Mid ATR	e [e]		o [o]
	Mid RTR	o [ɔ]		ɛ [ɛ]
	Low		a [a]	
Nasal	High	in [ï̃]		un [ũ]
	Low		an [ǣ̃]	

The phonetic transcription of the vowels is in square brackets and accompanied with the Standard Yorùbá orthography. Except in graphs and tables, the orthographic transcription is used throughout this paper. The low nasal vowel “an” is often pronounced as “on” (Bamgboṣe, 1966; Awóbùlúyì, 1978). Considering that the difference between “on” and “an” is phonetic, Pulleyblank (1988, p. 237) analyze the free variation between the vowels to be ‘a low-level phonetic effect.’ This phonetic distinction between “on” and “an” is not crucial to the goal of this paper. We now turn to how the tones and vowels are vocalized in Ịjálá and Ẹsà.

3. Basic description of Ịjálá and Ẹsà

Ịjálá and Ẹsà are some of the genres of oral poetry in Yorùbá culture. Most genres of Yorùbá oral poetry are associated with deities and ancestral devotion. Ịjálá is associated with Ọ̀gún, the Yorùbá deity of metallurgy and metal-related works (Babalọla, 1963; Ajuwon, 1977; Ọ̀látúnjí, 1979). Ọ̀gún is considered the patron of people who engage in metal works such as blacksmiths, goldsmiths, hunters and professional drivers. The devotees are obliged to pay homage to Ọ̀gún. One of the ways of paying homage to Ọ̀gún is through the chanting of oríkì, which is the embodiment of the eulogy and epithets about an entity, in this case about Ọ̀gún. The mode of performance of this oríkì in chant form is referred to as Ịjálá. In Yorùbá oral history, there are four hypotheses about the origin of Ịjálá. All the hypotheses point to the Yorùbá deity of metallurgy Ọ̀gún, but only two trace the origin of the vocal expression in Ịjálá to Ọ̀gún’s geriatric voice and his alcohol consumption (see Babalọla, 1963, p. 6–12).

Ẹsà, which is also known as Ịwí Egúngún, is associated with Egúngún creed of ancestral veneration (Olajubu, 1974; Adedeji, 1978). Periodically, the ancestral spirits physically manifest as Egúngún masquerade. To pay homage to the spirits of the departed ancestors, the devotees chant praises in Ẹsà poetic mode.

Unlike most genres of music in the culture, most genres of Yorùbá oral poetry are not danceable and may not be

accompanied with a drum performance. However, instrumental or vocal music can occur during the intermissions of the poetry performance (Babalola, 1963; Ajuwon, 1977; Adedeji, 1978; Fámúle, 2018). The instrumental music might involve the representation of Yorùbá phrases with a talking drum (Euba, 1990; Villepastour, 2010; Akinbo, 2019, 2021). Depending on the genre or the mood of the performer, the oral performance may be closer to speech or music, thus they are considered semi-musical verbal arts (Babalola, 1963; Ọlátúnjí, 1979; Ọ̀gúndẹ̀jì, 1991). Similar to Yorùbá, the genres of oral poetry in other African societies are also semi-musical (Uzochukwu, 1981; Boadi, 1989; Okpewho, 1992; Finnegan, 2007, 2012; Purvis, 2009).

Studies suggest that Yorùbá oral poetry can be identified based on the contents of the poem, the identity of the performer and vocal expression (Gbadamosi and Beier, 1959; Ọlátúnjí, 1979). The traditional contents of Ịjálá and Ẹ̀sà are eulogy and epithets, but the contents may include historical events, personal eulogy of hunters and non-hunters, social commentaries, humor and all aspects of human existence (see Babalola, 1963; Ọlátúnjí, 1979; Idamoyibo, 2006). Most importantly, the texts of a specific genre (e.g., Ịjálá) can be used for other genres of Yorùbá oral poetry (e.g., Ẹ̀sà) (Gbadamosi and Beier, 1959; Ọlátúnjí, 1979). As a result of this, textual contents are not reliable in distinguishing various genres of Yorùbá oral poetry. For example, the popular Ịjálá chanter, Ọ̀gúndáre Fọ́yánmu, is widely known for incorporating contemporary socio-political issues in his poems (Olaniyan, 2013). A noteworthy example is the Ịjálá chant of Fọ́yánmu about the historic Nigerian tax war of 1969 (see Adeniran, 1974)¹. Other examples come from the syncretic practices of Yorùbá Christians and Muslims. Although Ịjálá is traditionally associated with Ọ̀gún, Yorùbá Christians and Muslims often incorporate Ịjálá chants (and analogously other genres of Yorùbá verbal art) in their religious worships (Idamoyibo, 2006, 2011; Ajibade, 2007; Ọlátúnjí, 2012; Dada, 2014). Thus, the content of a poem and the identity of the performer are not reliable in identifying genres of Yorùbá oral poetry.

There is consensus that various genres of Yorùbá oral poetry are best distinguished or classified based on vocal expression (e.g., Ọlátúnjí, 1979; Yai, 1989, etc.). For example, the contents of Ịjálá are always chanted in vibrato (Babalola, 1963). Ẹ̀sà does not involve vibrato like Ịjálá, but a pattern of vowel insertion and lengthening (Olajubu, 1974; Adedeji, 1978). Regardless of the subject matter, a triply long vowel [ooo] at the end of the first poetic line is a recurrent characteristics of Ẹ̀sà (Olajubu, 1974; Olajubu and Ojo, 1977; Adedeji, 1978). When Christian and Islamic musicians incorporate Yorùbá oral poetry into their religious worships, the genres of oral poetry are recognized, not through the contents of the poem nor the identity of chanters,

¹ An audio of the Ịjálá poem by Fọ́yánmu can be found in this link: <https://www.youtube.com/watch?v=KzTBR7VJknQ>.

but the distinctive vocal expression which is associated with the chant.

Ịjálá shares the same vocal expression with Ịrèmòjé, which is a funeral dirge for hunters (Babalola, 1963; Ajuwon, 1977). Although Ịjálá can be adapted to suit any content, Ịrèmòjé is restricted to funeral rites. As a result of this, there is an on-going debate as to whether Ịrèmòjé is a distinct genre or a sub-genre of Ịjálá (see Babalola, 1963; Ajuwon, 1977; Olajubu, 1984; Idamoyibo, 2006). The features of vocal expressions in Ịjálá have also been described in terms of rhythm (Babalola, 1963; Ọlabimtan, 1977), but we do not consider rhythm in this work. Previous studies suggests that vocal expression in Yorùbá oral poetry involves stress given that it involves loudness and pitch raising (Siertsema, 1959; Babalola, 1963; Ọlabimtan, 1977; Ọlátúnjí, 1979). The vocal expression in Yorùbá oral poetry and stress have loudness and pitch raising in common, but the phonetic properties are not as a result of stress in the oral poetry. Unlike stress which involves a syllable being prominent than the other in a word (Lieberman and Prince, 1977; Halle and Idsardi, 1995; Kager, 2007), all the syllables of the words in Yorùbá oral poetry are produced with loudness and pitch raising (Babalola, 1963). Most importantly, Yorùbá is a tone language, not a stress-timed language (Akinlabi, 1985; Pulleyblank, 1986; Kenstowicz, 2006).

In this work, we investigate the phonetic correlates of vocal expression in Ịjálá and Ẹ̀sà. Yorùbá, the textual bases of the poem, is a tone language, which means pitch contrasts bring about lexical distinctions. Considering that the melody of verbal arts such as chanting depends on pitch contour, we also investigate how poetic melodies interact with the linguistic demand of tone contrast. To answer the questions, we conducted an experiment. The details of the experiment are presented in the next section.

4. Methodology

4.1. Stimuli, participant and procedure

The stimulus in this work is an original poem which was composed by the third author of this paper. As shown in (3), the poem is written in Standard Yorùbá orthography. Oral performance is usually from memory, so in order to make it easier for the consultant to memorize, we selected the oríkí for its brevity. By selecting an original poem instead of a widely known traditional poem as the stimulus, we were able to control for the effect of content and familiarity.

One male native speaker of Standard Yorùbá was voluntarily recruited for this study. This consultant (age 28) was a fourth-year undergraduate of the Yorùbá study program at the Department of Linguistics and African Languages, University of Ibadan. The consultant had trainings in chanting various genres of Yorùbá oral poetry, including Ịjálá and Ẹ̀sà. A week

before he was scheduled for a recording session, the consultant memorized the oríkí that was composed for the study. At the recording session, he was asked to recite the poem six times in normal speech mode. After reciting the poem in speech mode, he chanted the poem six times in Ìjálá and six times in Èṣà.

The renditions of the poem in speech mode and Ìjálá mode were recorded in a quiet room at the sampling rate of 48.1 kHz in wav format. Following Babalola (1963, p. 121), each stretch “of utterances after a breath pause” is grouped as a line of the poem. In line with the observation in Qlátúnjì (1979), the utterances within each pair of breath pauses form a meaningful whole. Based on the chanting of the poem in Ìjálá and Èṣà, the text of the poem were grouped into four lines.

- (3) An original Yorùbá poem
- Line 1 Adédùntán Àbèjé ọmọ Bábàlọjà
“Adeduntan, the child of Babaloja”
- Line 2 eyinjúu Ọmọladùn bariọlá ọmọba Lépolóyìn
“the eyeball of Omoladun, the honorable princess of Lepoloyin”
- Line 3 tẹ́ẹ́rẹ́ gbajó ọmọ̀dò̀dò̀ àgbà, ìdílẹ̀kẹ̀ ẹ̀lẹ̀rìn-ìn ẹ̀yẹ
“suitably slim for dance, a wise child with a bead-befitting waist”
- Line 4 dúdú wù mí, dúdú dá mí l’órùn máa wolẹ
máa rọra olówó tí n’owó Ẹ́áánú
“(your) blackness is alluring, walk cautiously (you) benevolent rich”

The tones and vowels of the text in speech and poetic modes were manually annotated in Praat (Boersma, 2001). For the three tones in the language (i.e., H, L, and M), F0(Hz) values of the pitch contour were extracted at 25%, 50%, and 75% intervals. To replicate the pitch contours as they appear in Praat windows for data visualization and analysis, each tone is labeled in a serial order, in this case T1.1, T1.2, T1.3...T2.1, etc, as shown in Figure 1. For each serially labeled tone, F0(Hz) values were extracted at twelve intervals. We extracted F0 values, intensity, formant and three spectral measurements (namely CPP, Energy below 500 Hz and Hammarberg) from the annotations, using the Praat scripts written by Riebold (2013) and Xu (2013). Using the script tremor.praat (Brückl, 2021), we measured vibrato rate (rate of frequency tremor). The praat script only works on segments that are >3 s long, but the duration of all the vowels vocalized with vibrato is <1 s. To make each vibrato vocalization at least 3 s long, each of the vibrato vowels was sextupled by itself. It is from the sextupled form that we extracted vibrato rate. See Brückl (2021), Riebold (2013), and Xu (2013) for more details on the scripts.

In the next section, we discuss the motivation for each of the acoustic measurements that were utilized in this work. Our data and the R code of our statistical analysis are in the Supplementary material that is attached to this article.

4.2. Motivation for acoustic measurements in this work

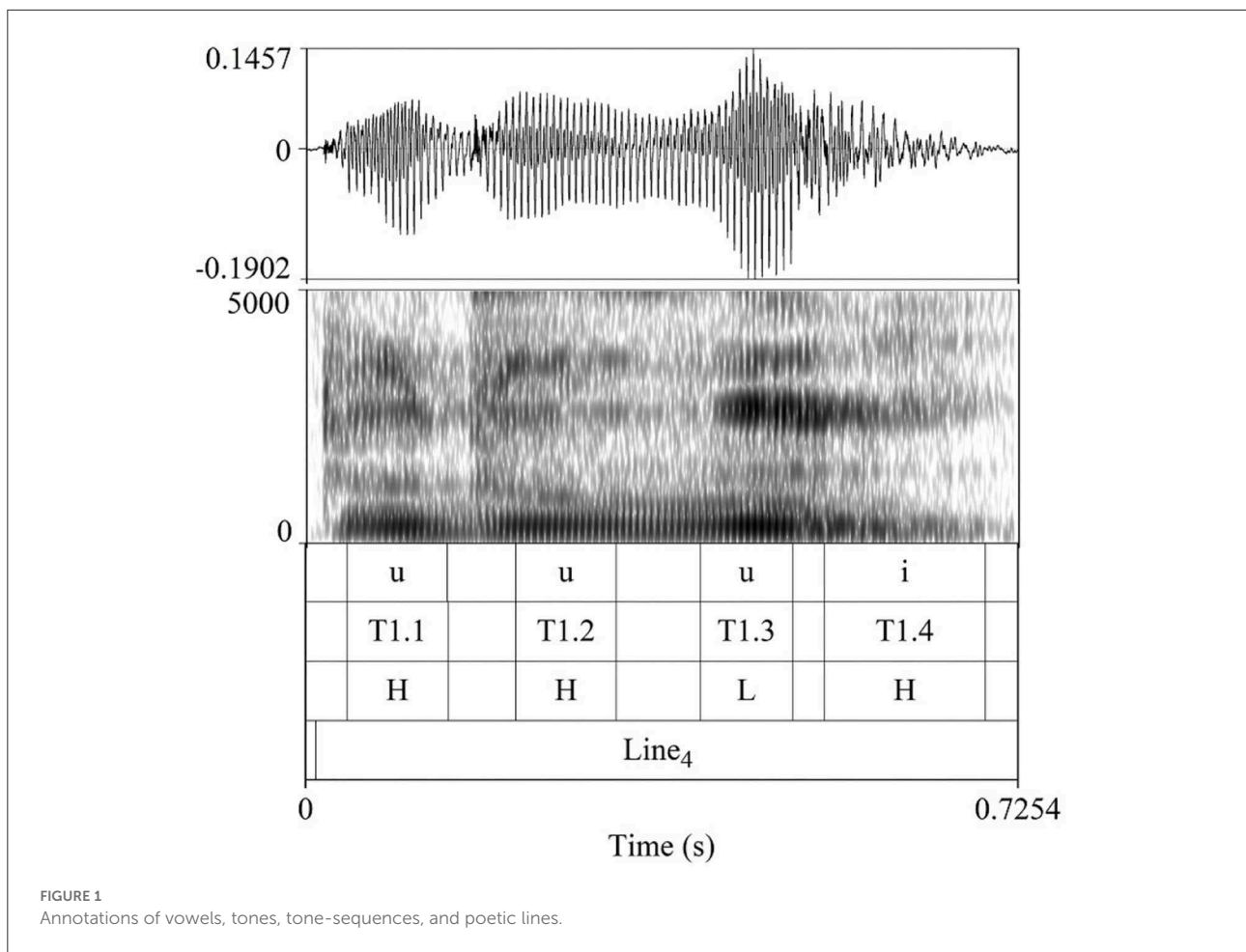
Nine acoustic parameters were measured for the annotated vowels in order to detect the acoustic correlates of vocal expression. The parameters are fundamental frequency, intensity, cepstral peak prominence, energy below 500 Hz, Hammarberg index, formant 1, formant 2, duration and frequency tremor. The parameters were selected based on the description of Ìjálá and Èṣà in previous studies. In this section, each of these parameters are described.

4.2.1. Fundamental frequency and intensity

We start the discussion with fundamental frequency (F0) which primarily depends on the vibration rate of the vocal cords. F0(Hz) is measured in hertz (Hz). The perceptual correlate of F0 is pitch (Ladefoged and Johnson, 2015). Pitch contrasts that bring about lexical or grammatical meaning distinctions are called tone (Yip, 2002; Hyman, 2018). As shown in Section 2, Yorùbá has three tones. Given that the vocal expressions in Yorùbá oral poetry involve pitch raising, it is crucial to investigate how pitch contours of speech melodies are mapped to poetic melodies. For this reason, we extracted the F0(Hz) values of the tones in speech and poetic modes. To capture the full-time interval of the pitch contours, we extracted F0(Hz) values at twelve intervals for each tone. Recall that increased loudness is also a property of vocal expression in Yorùbá oral poetry. The perceptual correlate of intensity is loudness, but the relationship is not linear. Consequently, we also measured the intensity of all vowels in speech and poetic modes.

4.2.2. Cepstral peak prominence: CPP (dB)

Cepstral peak prominence (CPP) is the difference between the maximum cepstral peak value occurring within the boundaries of the expected phonational quefrequencies and the corresponding value on the regression line fitted on the cepstrum (Hillenbrand et al., 1994; Hillenbrand and Houde, 1996). The degrees of glottal closure and vocal fold tension directly corresponds to the values of CPP(dB) (Kim, 1970; Iverson and Salmons, 1995; Inwald et al., 2011). Given that the glottal closure reduces the level of noise in vocal signal, CPP(dB) measures the level of noise in a vocal signal: the higher the noise, the lower the value of CPP and vice versa. In languages where the degree of glottal opening determines breathiness and aspiration, CPP(dB) is a reliable acoustic parameter for breathy and aspirated sounds (Hillenbrand et al., 1994; Blankenship, 2002; Esposito and Khan, 2012; Khan, 2012; Seyfarth and Garellek, 2018; Berkson, 2019). CPP(dB) was originally developed for measuring breathiness (Hillenbrand et al., 1994), but its usage has been extended to the evaluation of dysphonia. Studies show that the severity of dysphonia correlates with lower CPP values



when compared to normal voice (Hillenbrand and Houde, 1996; Heman-Ackah et al., 2002, 2003, 2014; Awan and Roy, 2005; Awan et al., 2009; Fraile and Godino-Llorente, 2014, etc.). As reported in Wolfe and Martin (1997), CPP(dB) values in hoarse and breathy voice are lower when compared to strained voice. As a result of the findings in various studies, the American Speech-Language-Hearing Association (ASHA) recommends CPP(dB) as a tool for “measuring the overall level of noise in the vocal signal” and as “a general measure of dysphonia” (Patel et al., 2018).

The use of CPP(dB) values has also been extended to the evaluation of effortful speech production and emotive speech. For example, increased loudness and sustained vowel production in effortful speech production correlate with higher CPP(dB) values when compared to normal speech production (Rosenthal et al., 2014; McKenna and Stepp, 2018; Phadke et al., 2020). Given that chanting involves increased loudness and high vocal demand, we could understand the vocal features of the poetic modes by measuring CPP(dB) values. Considering that nasality can decrease CPP(dB) values (see Madill et al., 2019), we only extracted CPP(dB) values for oral vowels.

4.2.3. Energy below 500 Hz

Another acoustic parameter which is used in this work is the proportion of spectral Energy below 500 Hz (dB). This measurement is often used for evaluating vocal quality in affective speech (Tolkmitt et al., 1982; Johnstone et al., 2005; Scherer et al., 2017). Low values of the Energy below 500 Hz are associated with the tensioning of vocal cords (Tolkmitt et al., 1982; Scherer et al., 2002, 2017; Johnstone et al., 2005). The values of Energy below 500 Hz (dB) were only extracted for oral vowels.

4.2.4. Hammarberg index (dB)

We extracted the values of Hammarberg index for evaluating vocal expression. The Hammarberg index is defined as the difference between the energy maximum in the 0–2,000 Hz frequency band and in the 2,000–5,000 Hz band (Hammarberg et al., 1980). Studies suggest that increase in loudness and F0(Hz) correlates with lower values of Hammarberg index (Scherer et al., 2017; Hakanpää et al., 2021; Sundberg et al., 2021). As an addition measurement for pitch raising and increased loudness,

we measured Hammarberg index for all oral vowels in speech and poetic modes.

4.2.5. Formant frequencies

Formants are defined as “a resonating frequency of the air in the vocal tract” (Ladefoged and Johnson, 2015, p. 315). The first two formants, namely formant 1 (F1) and formant 2 (F2) are important in determining the quality of vowels. Specifically, F1(Hz) is mostly determined by vowel height and F2(Hz) is determined by vowel frontness or backness. The values of F1 increases in loud and effortful speech and verbal arts, but the values of F2 is not consistent under the same condition (Huber et al., 1999; Traunmüller and Eriksson, 2000; Huber and Chandrasekaran, 2007; Koenig and Fuchs, 2019, etc.). To understand the effect of vocal expression on the acoustics of vowels, F1(Hz) and F2(Hz) values were extracted for all the oral vowels.

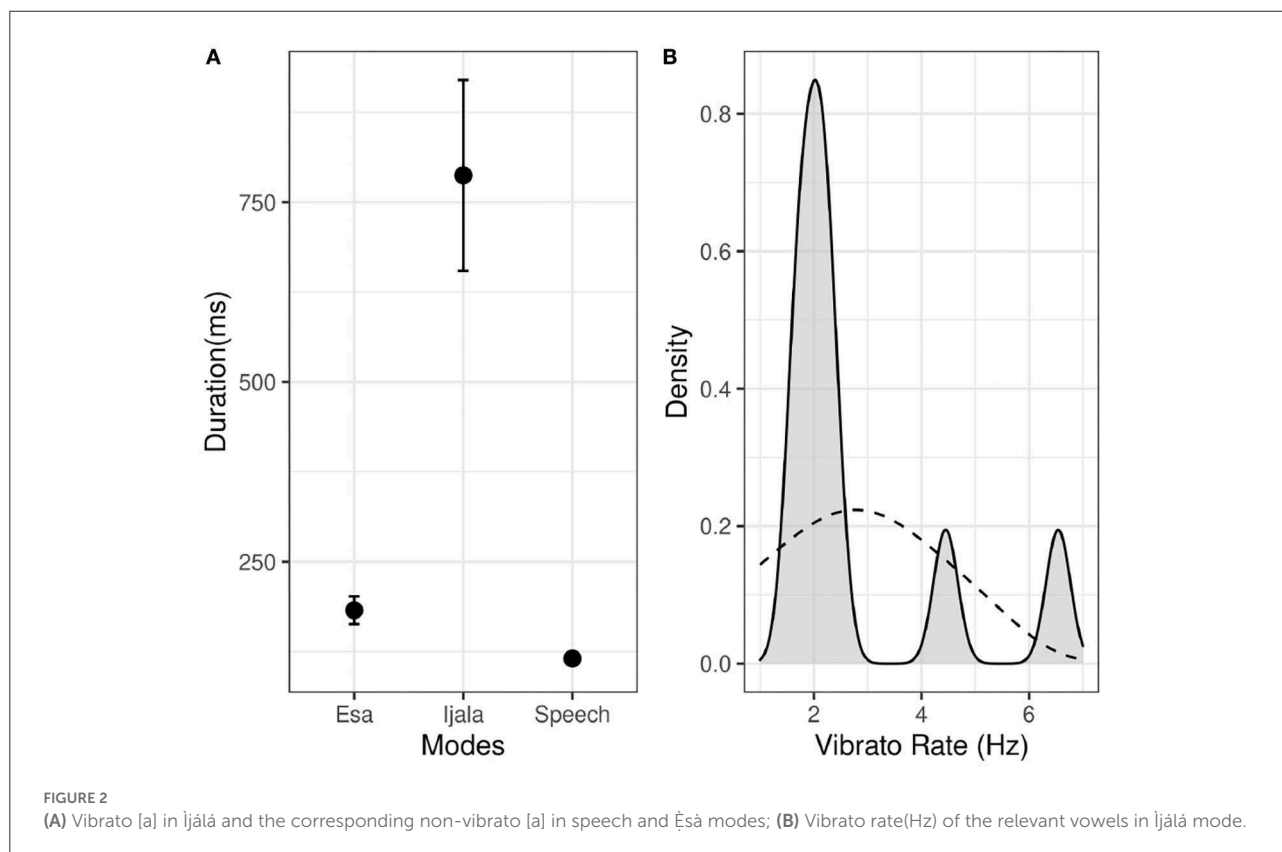
4.2.6. Vibrato rate (frequency tremor)

We also measured the rate of frequency modulation or tremor. When frequency modulation occurs as a result of alcohol withdrawal syndrome (Koller et al., 1985; Anouti and Koller, 1995), aging (Gregory et al., 2012; Martins et al., 2015),

or neurological disorders that cause involuntary movement of muscles in the throat, larynx (voice box), and vocal cords, it is called vocal tremor (Deuschl et al., 1998; Hlavnička et al., 2020). When used intentionally in singing, frequency tremor is called vibrato (Seashore, 1938; Dromey et al., 2003; Nix et al., 2016). The typical values of vibrato rate range from 4 to 7 Hz (Seashore, 1938; Dromey et al., 2003; Nix et al., 2016). In neurological diseases, vocal tremor frequencies are categorized into slow (<4 Hz), intermediate (4–7 Hz) or rapid (>7 Hz) (Deuschl et al., 1998; Charles et al., 1999). The slow tremor frequencies are prominent in all neurological disorders, but the intermediate and rapid tremor are mostly found in a subset of neurological disorders (Deuschl et al., 1998; Hlavnička et al., 2020).

4.3. Statistical analysis

The linear-mixed effect model was fitted to each acoustic parameters for each vowel and tone, to determine whether speech and poetic modes have a significant effect. In this case, the fixed effect is the modes (i.e., speech, *ljálá* and *Ěsà*), and the random effect is each iteration of the poems in all modes. For the tones, the random effect is the tone-bearing unit, in this case the vowels. This was done using the package “lme4” in R (Bates et al., 2014). We ran *post-hoc* pairwise comparisons for the mixed



effect model using the package “emmeans” (Lenth and Lenth, 2018).

To calculate the correspondence between pitch trajectories of speech and poetic modes, we used Pearson correlation coefficient R which measures the strength and direction of a linear relationship between two variables. The value of R is always between +1 and -1. The closer the value of R is to +1, the stronger the positive relationship between the two variables. However, the closer the value of R is to -1, the stronger the negative relationship between the two variables. If the value of R is 0, it means there is no relationship between the two variables (see Rumsey, 2009, for a basic description this statistical measurement). A regression line in a scatter plot describes the strength and direction of the linear relationship between the variables under consideration.

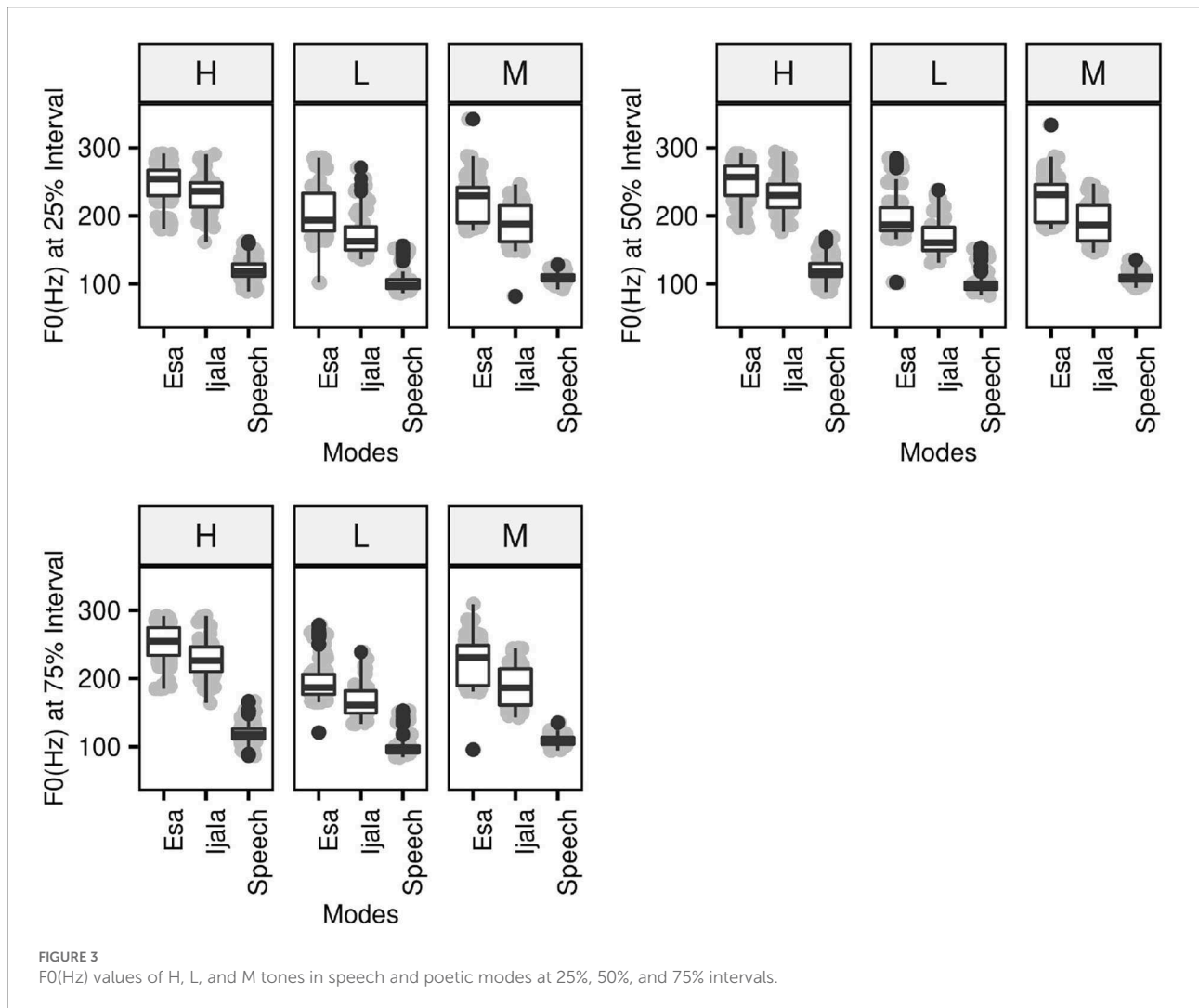
The null hypothesis is that there is no difference between speech and poetic modes for all the acoustic parameters. If the $p \leq 0.05$, there is a statistically significant effect of speech or poetic

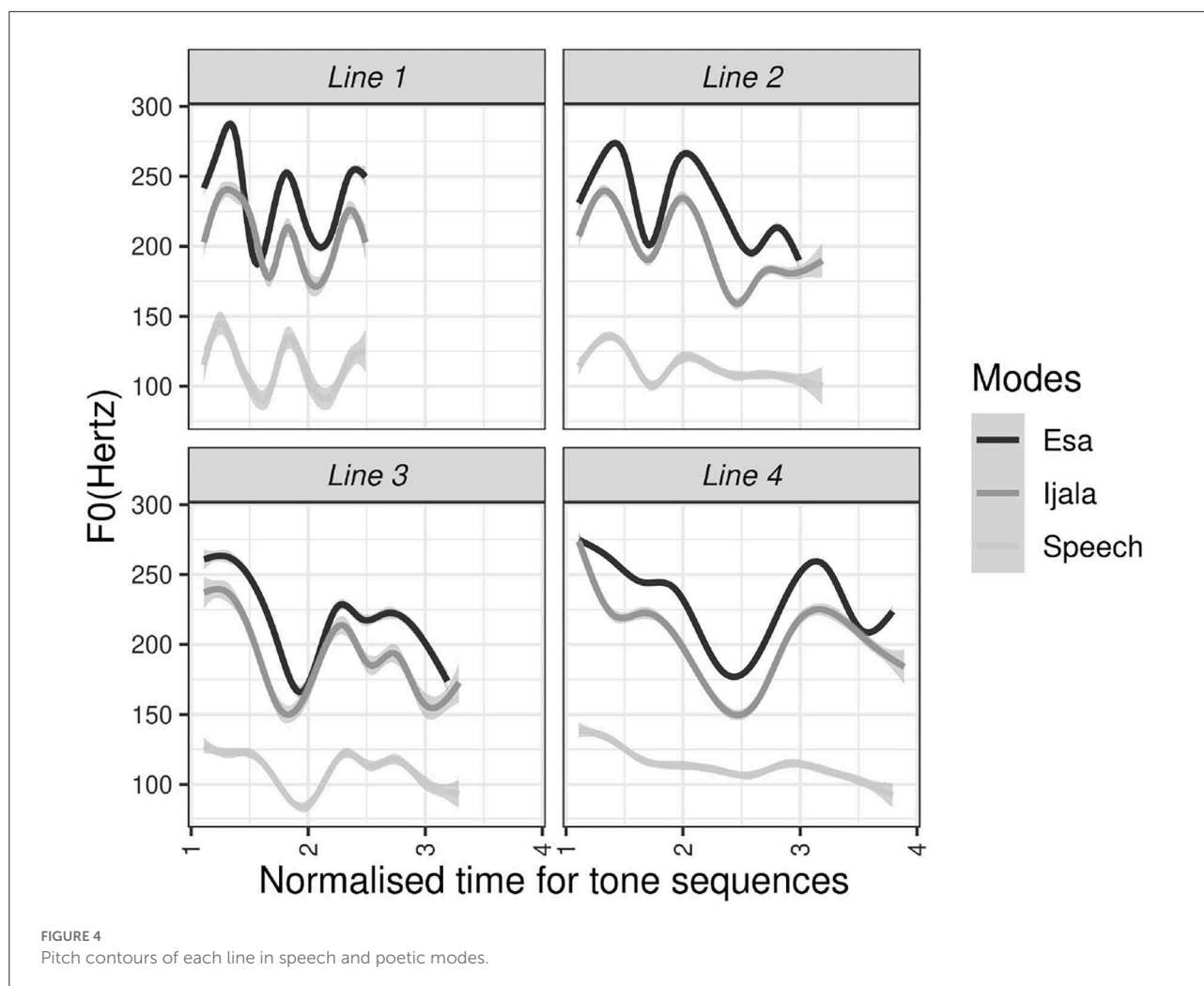
modes for the acoustic parameters. Therefore, we have a strong evidence against the null hypothesis. A $p > 0.05$ indicates weak evidence against null hypothesis (Rumsey, 2009; Wasserstein and Lazar, 2016). In the next section, we present the results of the acoustic analysis.

5. Results

We discuss the phonological attribute of vocal expression, before turning to the results of the acoustic analysis. In Èsà, the triply long vowel [ooo] is inserted at the beginning of the first word in line 1 even though the text does not have such vowel. If we recall that this is a recurrent attribute of Èsà, we can say the long vowel is an attribute of vocal expression in Èsà.

We now turn to the results of the acoustic analysis. One syllable in each of the first three lines were lengthened and





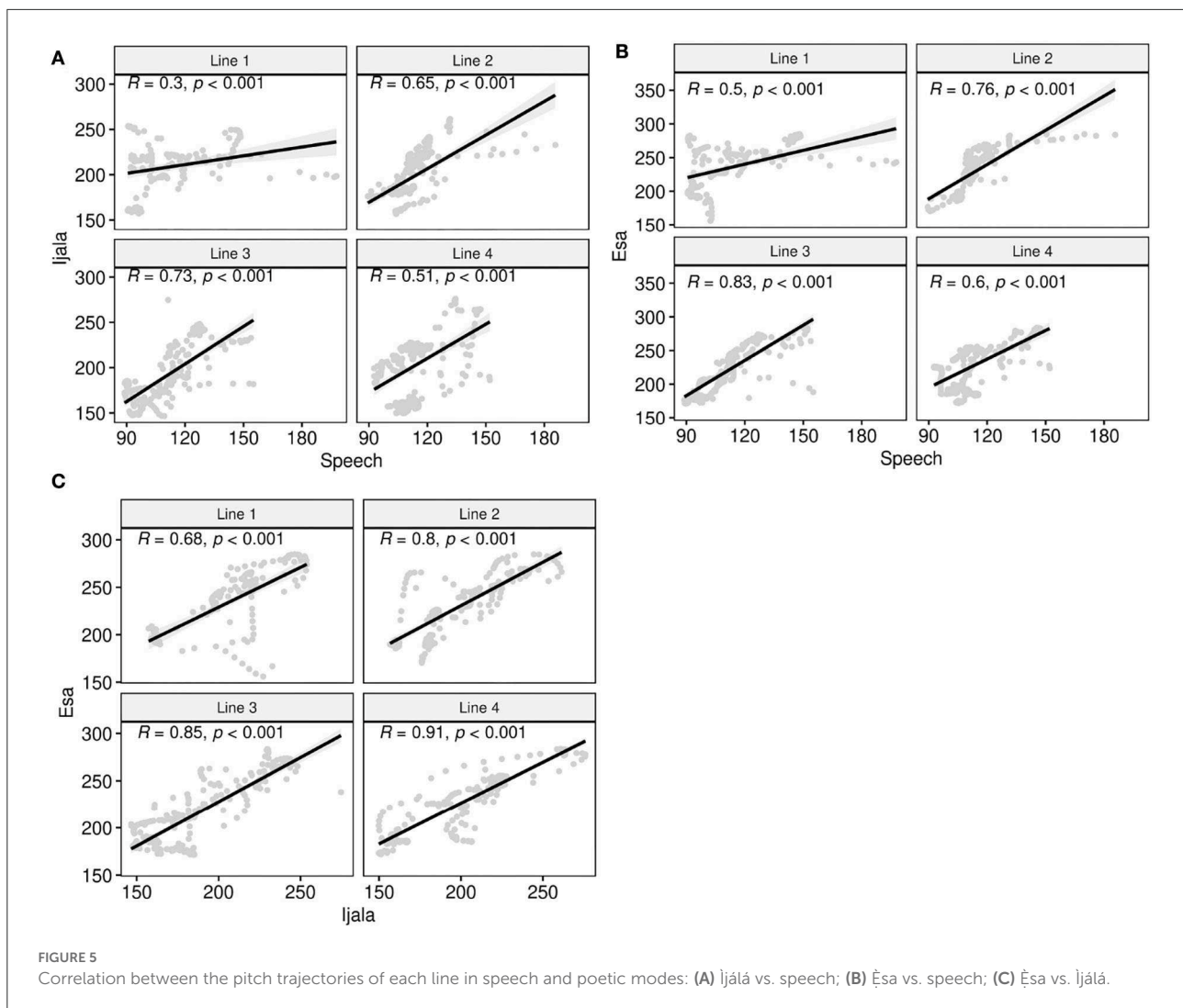
vocalized with vibrato in the Ìjálá mode, but at no point was Èsà vocalized with vibrato. All the syllables targeted for vibrato (except for one) are in the range of the last and penultimate word in each poetic line. The syllable that were consistently targeted for vibrato in each iteration of the poem contains the sequence [ba], but the syllables with the vowel [ɪ, ɛ] were variably vocalized with vibrato. Considering that the other vowels were not consistently targeted for vibrato, we only measured the duration of the vowel that was consistently targeted for vibrato, as shown in Figure 2A. The vowel targeted for vibrato in the Ìjálá mode is significantly longer than the corresponding vowel in Èsà ($p < 0.001$), which in turn is longer than the corresponding vowel in speech mode. However, the distinction between Èsà and speech modes for the vibrato [a] is not statistically significant.

All the vowels that were produced with vibrato in Ìjálá have a vibrato rate in the range 1.6–2 Hz, as shown in Figure 2B. However, the vowels [ɛ] and [a] have the vibrato rate of 4.45 and 6.5 (Hz) respectively in one of

their repetitions. Thus, the variation cannot be attributed to vowel-type. Considering that the vibrato rate of 4.45 (Hz) and 6.5 (Hz) are only found in two tokens, they are considered outliers.

Compared to Ìjálá, F0 values at 25%, 50%, and 75% intervals are higher in Èsà for the three tones, as shown in Figure 3. Also at every interval, the values of F0(Hz) for each tone are higher in the poetic modes than the speech mode. The vocalization modes have a significant effect on the values of F0(Hz) for speech, Ìjálá and Èsà comparisons ($p < 0.001$). This shows that the vocal expression targets all the tones in the language. The results also show that the three tones in the language have distinct F0(Hz) values regardless of the vocalization mode.

The pitch trajectory of the tone sequences in each poetic line is compared in Figure 4. In the figure, the y-axis contains the acoustic measurement of pitch contour in F0(Hz), and the x-axis contains the proportional time of tone sequences. The dark line is for the pitch contours of Èsà, the dark gray for Ìjálá and the light gray line is for the pitch contour of speech. There are four



panels in the graph, where each panel is for the sequence of tones in each poetic line.

As shown in Figure 4, the values of F0(Hz) are higher in poetic modes when compared to the speech mode. The values of F0(Hz) are higher in Èsà than lǐjálá. Figure 4 also shows that the pitch trajectory of each poetic line in speech mode is similar to that of the corresponding poetic line in Èsà and lǐjálá.

To check the degree of similarity between the pitch contours of speech and poetic modes, we applied Pearson correlation coefficient to the pitch contour. To investigate whether linear relationship between speech and poetic modes varies based on poetic lines, the correlation coefficient are applied to each of the four poetic lines for every speech, lǐjálá and Èsà comparisons. The results are shown in Figure 5.

Figure 5 shows that there are statistically significant positive correlations between the trajectories of speech and poetic melodies ($p < 0.001$), but the degree of correlation varies by poetic lines and genres. Comparing the lines in Figures 5A,B, we

see that the correlation between the pitch contours of Èsà and speech is higher when compared to the correlation between the pitch contours of lǐjálá and speech. Figure 5 also shows that the pitch contours of lǐjálá and Èsà are closer than they are to the pitch contours of speech.

Similar to the F0(Hz) values, the intensity of oral vowels is significantly higher in poetic modes than in speech mode ($p < 0.001$), as shown in Figure 6. However, the distinction between the intensity Èsà and lǐjálá varies depending on vowel-type. As shown in Figure 6, the intensity is higher in Èsà than lǐjálá for all vowels, except for the vowel [u]. The difference between the intensity of Èsà and lǐjálá is only significant ($p \leq 0.004$) for the vowels [o, ɔ, a]. The values of the Energy below 500 Hz(dB) are significantly higher in speech than poetic modes for all vowels ($p \leq 0.004$), except the vowels [e, o, ɔ]. The distinction between the Energy below 500 Hz(dB) of Èsà and lǐjálá is only significant for [o, ɔ]. We now turn to the results of CPP(dB) and Hammarberg index(dB), which are presented Figure 7.

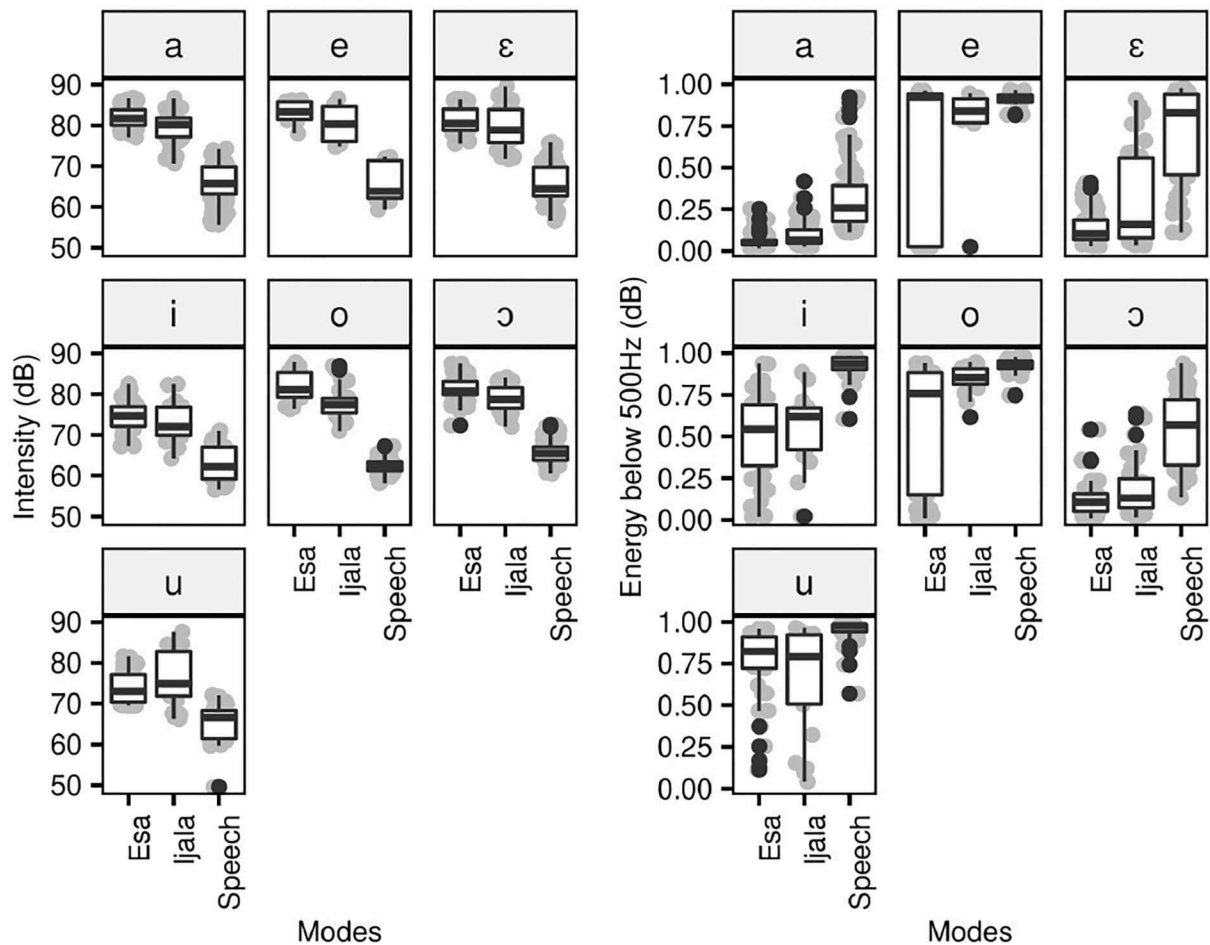


FIGURE 6 Intensity and energy below 500(Hz) of oral vowels in poetic and speech modes.

The results of the statistical analysis indicate that the mean values of CPP(dB) are significantly higher in poetic modes when compared to speech mode ($p < 0.001$). However, the difference between the CPP(dB) values of Èsà and Ìjálá is not significant. The values of Hammarberg index are significantly lower in poetic modes than speech mode ($p < 0.001$). The distinction between the Hammarberg index(dB) values of Èsà and Ìjálá is only statistically significant for the vowels [i, u, ɔ] ($p \leq 0.013$). The results of the formant values are presented in Figure 8.

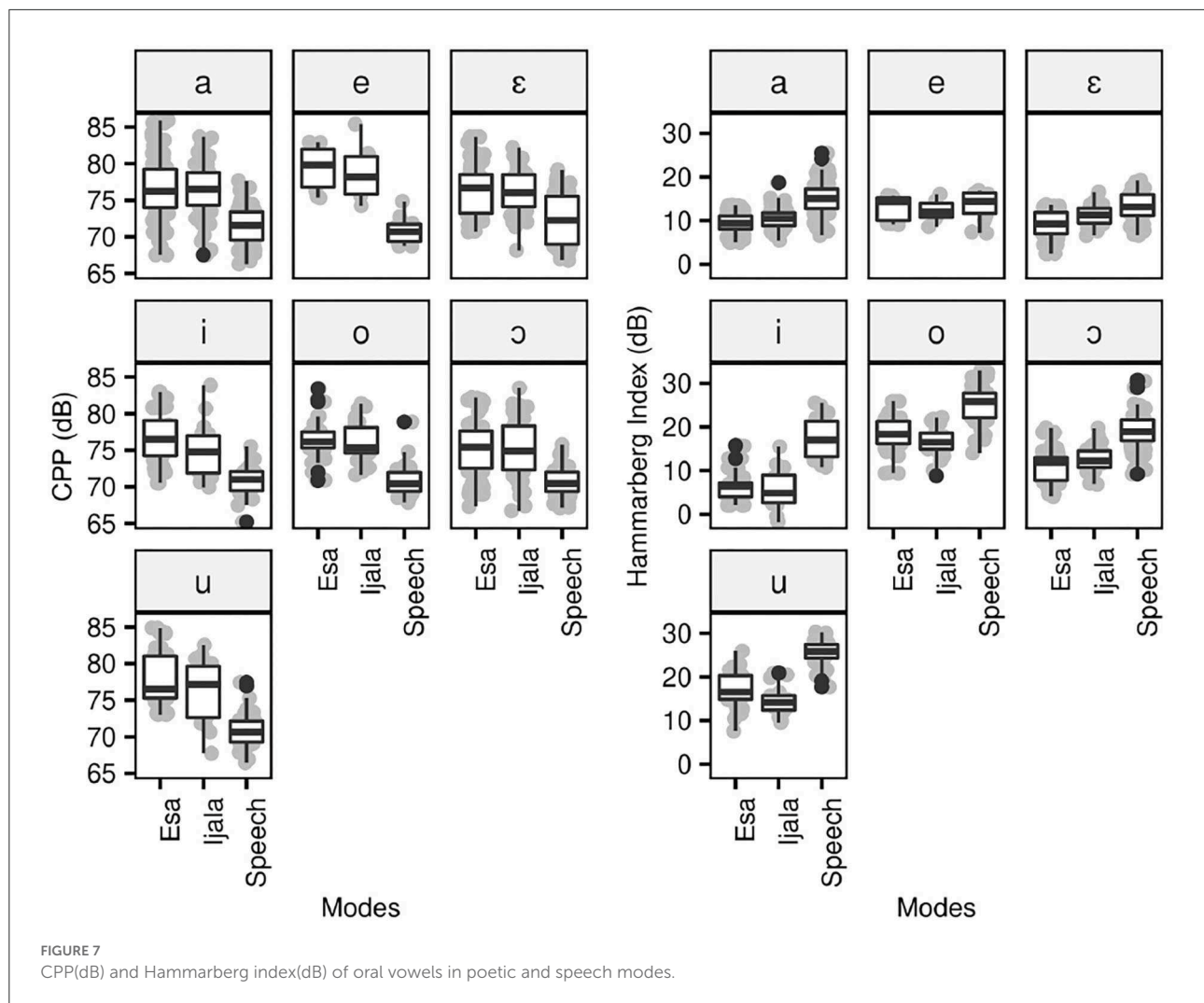
There is an effect of vocal expression on vowel formants. As shown in Figure 8, the values of F1 for all oral vowels are higher in poetic modes than speech mode. For the values of F1(Hz), the distinction between poetic modes and speech mode is significant for all vowels ($p \leq 0.019$), except the vowel [u]. The graph in Figure 8 also shows that, for all vowel except [u], the values of F1 are slightly higher in Èsà than Ìjálá. However, the F1 distinction between Èsà than Ìjálá is only significant ($p \leq 0.026$) for the vowels [o, ɔ, a]. We now turn the values of F2(Hz). There is no

obvious distinction between the values of F2(Hz) for speech, Èsà and Ìjálá, except for the vowels [e, ε] that have lower F2(Hz) in poetic modes. Even in this case, the distinction is only significant for the vowel [ε].

In the next section, we discuss the results and their implications for the analysis of vocal expression in Èsà and Ìjálá.

6. Discussion and conclusion

We set out in this work to understand the acoustic correlates of vocal expression in Ìjálá and Èsà, under an experimental condition. The results of our investigation show that Hammarberg index(dB), Energy below 500 Hz(dB) and F1(Hz) distinguish the speech mode from each of the poetic modes for some vowels but are not as reliable as vibrato, F0(Hz), CPP(dB) and intensity(dB) which distinguish the poetic modes from each of the poetic modes for all vowels. For Èsà vs. Ìjálá, Èsà



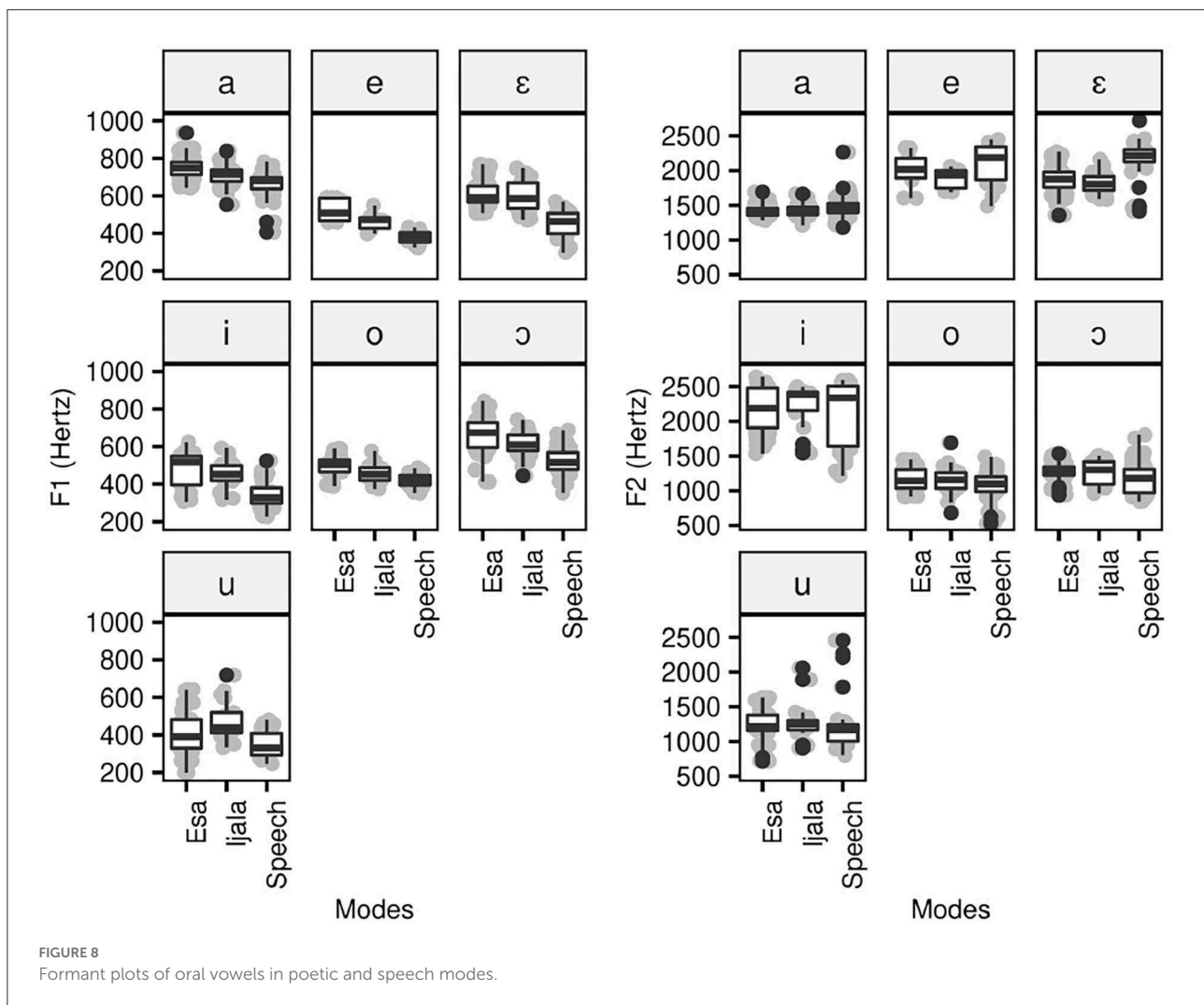
vs. speech and *Ījálá* vs. speech comparisons, The most reliable acoustic parameters are vibrato and F_0 (Hz), given that vibrato only features in *Ījálá* and that there is a significant effect of vocalization modes on F_0 (Hz), regardless of tone and poetic line. The results also show that there is a correspondence between the pitch trajectories of speech tones and poetic tunes, but the degree of correspondence varies by poetic lines and genres. Another distinctive feature that distinguishes *Èsà* from *Ījálá* is the epenthesis and lengthening of the vowel [o] in the first poetic line.

The acoustic correlates of vocal expressions in *Èsà* and *Ījálá* are consistent with increased vocal effort, given that higher values of F_1 (Hz), intensity and CPP are associated with increased vocal effort (e.g., [Jessen et al., 2005](#); [Rosenthal et al., 2014](#); [McKenna and Stepp, 2018](#)). The lower values of Energy below 500 Hz in poetic modes are also consistent with vocal tensing found in effortful speech. An increased vocal effort is expected as a feature of both genres, considering that

vocal performance in a large space requires high vocal effort ([Sundberg, 1977](#); [Beechey et al., 2018](#)) and that *Ījálá* and *Èsà* are typically performed to a large audience in an open space ([Babalola, 1963](#); [Adedeji, 1978](#); [Yai, 1989](#)). It is probably vocal effort that previous research mischaracterised as stress.

The range of vibrato rate (1.6–2 Hz) in *Ījálá* is atypical of the vibrato rate (4–7 Hz) in singing but consistent with vocal tremor as the historical origin of vocal expression in *Ījálá*. Although the range of the vibrato rate reported in this work is prominent in all neurological diseases, it is difficult to tell whether the vibrato in *Ījálá* historically developed from the vocal symptoms of alcohol withdrawal, aging or both. The pitch-height distinction between *Ījálá* and *Èsà* cannot be attributed to vibrato considering that the vibrato and non-vibrato sections of *Ījálá* have lower pitch height than *Èsà*, as shown in [Figure 4](#).

Another notable finding of this study is tone-tune mapping. Studies on tone-tune mapping indicate that song melodies in a tone language are not determined by language, but music



can accommodate language when it is musically feasible (Ho, 2006; Schellenberg, 2009, 2013; McPherson and Ryan, 2018). The results of our study is in line with the findings of studies on tone-tune mapping in singing, given that the correspondence relations between the pitch contours of speech tones and poetic tunes varies based on genres and poetic lines. As shown in the results of the correlation coefficient in Figure 5, the tune of Èsà is closer to speech-tone melody than the tune of Ìjálá. This indicates that Èsà is closer to speech than Ìjálá, in terms of tone-tune mapping. It remains to be seen whether this makes the chants of Èsà more intelligible than Ìjálá.

Studies on affective 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 (Banse and Scherer, 1996; Juslin and Laukka, 2003; Johnstone et al., 2005; Goudbeek and Scherer, 2010; Lindquist et al., 2016; Scherer, 2021). It remains to be seen whether the pitch raising and increased loudness in Ìjálá and Èsà are also associated with high level arousal such as excitement and happiness. Considering

that Ìrèmòjé is a dirge with similar vocal expressions as Ìjálá, future research involving more participants should compare the acoustic cues of vocal expression in Ìjálá and Ìrèmòjé.

The major limitation of this work is that it is based on data from one participant. As a result of this, we cannot tell whether the acoustic correlates of vocal expression in this work are specific to the single participant or applies to other Yorùbá chanters. Thus, future research should replicate the present study on a larger population of Yorùbá chanters. Another limitation of this research is that we did not specifically look at the effect of vibrato on each lexical tone. To the best of our knowledge, the effect of vibrato on tone has not been studied, future research on singing or chanting in a tone language should investigate the interaction of tone and vibrato.

In sum, our study supports the observation in previous studies that vocal expressions, such as pitch raising, vowel epenthesis and lengthening, distinguish Ìjálá, Èsà and speech. Contrary to the previous impressionistic observations, increased loudness as vocal expression does not distinguish Ìjálá from

Èṣà but the poetic modes from speech. In addition, we have shown that the vocal expression in Yorùbá oral poetry might be attributed to high vocal effort. Our analysis of tone-tune mapping in the poetic modes indicates that the poetic tunes correspond to the melody of speech tones, but the degree of correspondence varies based on poetic lines and genres. In addition to the analytical importance, the present study also supports vocal tremor as the historical origin of vocal expression in Ìjálá. It is important to note that the properties of vocal expression reported in this work were not possible to capture through older impressionistic observation methods. Given that properties of vocal expression in oral poetry are better captured in phonetic terms, we strongly recommend the instruments of phonetic science as valuable tools for the study of African verbal arts.

Author's note

This research was carried out when the first author was a student working under an SSHRC insight grant (435-2016-0369) awarded to Douglas Pulleyblank at the University of British Columbia. For comments and suggestions on various aspects of this work, we thank Douglas Pulleyblank, Rose-Marie Dechaine, Adélékè Adéèkó, Tolu Odebunmi, Akosua Addo, and the reviewers. Errors of fact or explanation are our own responsibility.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation

References

- Adedeji, J. (1978). The poetry of the Yoruba masque theatre. *Afr. Arts* 11, 62–100. doi: 10.2307/3335415
- Adeniran, T. (1974). The dynamics of peasant revolt: A conceptual analysis of the Agbekoya Parapo uprising in the western state of Nigeria. *J. Black Stud.* 4, 363–375. doi: 10.1177/002193477400400401
- Ajibade, G. O. (2007). New wine in old cups: Postcolonial performance of Christian music in Yorùbá land. *Stud. World Christian.* 13, 105–126. doi: 10.3366/swc.2007.13.2.105
- Ajuwon, B. (1977). *Funeral Dirges of Yoruba Hunters* (Ph.D. Thesis). Indiana University.
- Akinbo, S. K. (2019). Representation of Yorùbá tones by a talking drum: an acoustic analysis. *Linguist. Lang. Afr.* 5, 11–23. doi: 10.4000/lla.347

and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

SKA designed the experiment, analyzed the data, write up the background, results, and discussion. OS collected the data, analyzed the data, and edited the manuscript. IBA set up the stimulus and write up the background. OA collected the data and edited the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2022.1029400/full#supplementary-material>

- Akinbo, S. K. (2021). The language of gárgan, a Yorùbá talking drum. *Front. Commun.* 6, 650382. doi: 10.3389/fcomm.2021.650382

- Akinlabi, A. (1985). *Tonal Underspecification and Yoruba Tone* (Ph.D. Thesis). University of Ibadan.

- Anouti, A., and Koller, W. C. (1995). Tremor disorders, diagnosis and management. *Western J. Med.* 162, 510.

- Awan, S. N., and Roy, N. (2005). Acoustic prediction of voice type in women with functional dysphonia. *J. Voice* 19, 268–282. doi: 10.1016/j.jvoice.2004.03.005

- Awan, S. N., Roy, N., and Dromey, C. (2009). Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clin. Linguist. Phonet.* 23, 825–841. doi: 10.3109/02699200903242988

- Awobùlúyì, O. (1978). *Essentials of Yorùbá Grammar*. Ibadan: Oxford University Press Nigeria.
- Babalola, S. A. (1963). *The Content and Form of Yoruba Ijala* (Ph.D. Thesis). SOAS University of London.
- Bamgboṣe, A. (1965). *Yoruba Orthography: A Linguistic Appraisal with Suggestions for Reform*. Ibadan: Ibadan University Press.
- Bamgboṣe, A. (1966). *A Grammar of Yoruba, Vol. 5*. Cambridge: Cambridge University Press.
- Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614. doi: 10.1037/0022-3514.70.3.614
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1. doi: 10.18637/jss.v067.i01
- Beechey, T., Buchholz, J. M., and Keidser, G. (2018). Measuring communication difficulty through effortful speech production during conversation. *Speech Commun.* 100, 18–29. doi: 10.1016/j.specom.2018.04.007
- Berkson, K. H. (2019). Acoustic correlates of breathy sonorants in Marathi. *J. Phon.* 73, 70–90. doi: 10.1016/j.wocn.2018.12.006
- Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *J. Phon.* 30, 163–191. doi: 10.1006/jpho.2001.0155
- Blench, R. (2019). Niger-congo: An alternative view. Manuscript. Available online at: <http://www.rogerblench.info/Language/Niger-Congo/General/Niger-Congo%20an%20alternative%20view.pdf> (accessed January 10, 2021).
- Boadi, L. A. (1989). Praise poetry in Akan. *Research in African Literatures*, 20, 181–193.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott Int.* 5, 341–345.
- Brown, D. J. B. (1995). *Orality, Textuality and History: Issues in South African Oral Poetry and Performance* (Ph.D. Thesis). Rutgers University.
- Brückl, M. (2021). tremor3.05. praat-script for the computation of 18 measures of (vocal) tremor. doi: 10.13140/RG.2.2.13850.57287
- Charles, P. D., Esper, G. J., Davis, T. J., Maciunas, R. J., and Robertson, D. (1999). Classification of tremor and update on treatment. *Am. Fam. Physician* 59, 1565.
- Dada, A. O. (2014). Old wine in new bottle: elements of Yoruba culture in Aladura Christianity. *Black Theol.* 12, 19–32. doi: 10.1179/1476994813Z.00000000017
- Deuschl, G., Bain, P., Brin, M., and Committee, A. H. S. (1998). Consensus statement of the movement disorder society on tremor. *Mov. Disord.* 13, 2–23. doi: 10.1002/mds.870131303
- Dromey, C., Carter, N., and Hopkin, A. (2003). Vibrato rate adjustment. *J. Voice* 17, 168–178. doi: 10.1016/S0892-1997(03)00039-0
- Esposito, C. M., and Khan, S. U. D. (2012). Contrastive breathiness across consonants and vowels: a comparative study of Gujarati and White Hmong. *J. Int. Phon. Assoc.* 42, 123–143. doi: 10.1017/S0025100312000047
- Euba, A. (1990). *Yorùbá Drumming: The Dùndún Tradition*. Bayreuth: Bayreuth African Studies.
- Fámùle, O. (2018). Èdè àyàn: The language of Àyàn in Yorùbá art and ritual of Egúngún. *Yoruba Studies Review*, 2, 1–50.
- Finnegan, R. (2007). *The Oral and Beyond: Doing Things with Words in Africa*. London: James Currey; University of Chicago Press.
- Finnegan, R. (2012). *Oral Literature in Africa*. Cambridge, UK: Open Book Publishers.
- Fraile, R., and Godino-Llorente, J. I. (2014). Cepstral peak prominence: comprehensive analysis. *Biomed. Signal Process. Control* 14, 42–54. doi: 10.1016/j.bspc.2014.07.001
- Gbadamosi, B., and Beier, U. (1959). *Yoruba Poetry*. Ibadan: Government Printers.
- Goudbeek, M., and Scherer, K. (2010). Beyond arousal: valence and potency/control cues in the vocal expression of emotion. *J. Acoust. Soc. Am.* 128, 1322–1336. doi: 10.1121/1.3466853
- Gregory, N. D., Chandran, S., Lurie, D., and Sataloff, R. T. (2012). Voice disorders in the elderly. *J. Voice* 26, 254–258. doi: 10.1016/j.jvoice.2010.10.024
- Hakanpää, T., Waaramaa, T., and Laukkanen, A.-M. (2021). Training the vocal expression of emotions in singing: effects of including acoustic research-based elements in the regular singing training of acting students. *J. Voice*. doi: 10.1016/j.jvoice.2020.12.032
- Halle, M. and Idsardi, W. (1995). “General properties of stress and metrical structure,” in *The handbook of phonological theory*, ed J. Goldsmith, (Cambridge, MA: Blackwell), 403–443
- Hammarberg, B., Fritzell, B., Gaufin, J., Sundberg, J., and Wedin, L. (1980). Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngol.* 90, 441–451. doi: 10.3109/00016488009131746
- Heman-Ackah, Y. D., Michael, D. D., Baroody, M. M., Ostrowski, R., Hillenbrand, J., Heuer, R. J., et al. (2003). Cepstral peak prominence: a more reliable measure of dysphonia. *Ann. Otol. Rhinol. Laryngol.* 112, 324–333. doi: 10.1177/000348940311200406
- Heman-Ackah, Y. D., Michael, D. D., and Goding Jr, G. S. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *J. Voice* 16, 20–27. doi: 10.1016/S0892-1997(02)00067-X
- Heman-Ackah, Y. D., Sataloff, R. T., Laureyns, G., Lurie, D., Michael, D. D., Heuer, R., et al. (2014). Quantifying the cepstral peak prominence, a measure of dysphonia. *J. Voice* 28, 783–788. doi: 10.1016/j.jvoice.2014.05.005
- Hillenbrand, J., Cleveland, R. A., and Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *J. Speech Lang. Hear. Res.* 37, 769–778. doi: 10.1044/jshr.3704.769
- Hillenbrand, J., and Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *J. Speech Lang. Hear. Res.* 39, 311–321. doi: 10.1044/jshr.3902.311
- Hlavnička, J., Tykalová, T., Ulmanová, O., Dušek, P., Horáková, D., Růžička, E., et al. (2020). Characterizing vocal tremor in progressive neurological diseases via automated acoustic analyses. *Clin. Neurophysiol.* 131, 1155–1165. doi: 10.1016/j.clinph.2020.02.005
- Ho, W. S. V. (2006). “The tone-melody interface of popular songs written in tone languages,” in *Proceedings of the 9th International Conference on Music Perception and Cognition*, eds M. Baroni, A. R. Addressi, R. Caterina, and M. Costa (Bologna: Bononia University Press), 1414–1422.
- Huber, J. E., and Chandrasekaran, B. (2007). Effects of increasing sound pressure level on lip and jaw movement parameters and consistency in young adults. *J. Speech Lang. Hear. Res.* 2, 1368–1379. doi: 10.1044/1092-4388(2006)098
- Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., and Johnson, K. (1999). Formants of children, women, and men: the effects of vocal intensity variation. *J. Acoust. Soc. Am.* 106, 1532–1542. doi: 10.1121/1.427150
- Hyman, L. M. (2018). What tone teaches us about language. *Language* 94, 698–709. doi: 10.1353/lan.2018.0040
- Idamoyibo, A. A. (2006). *Yoruba Traditional Music in Christian Worship: A Case Study of Ijala Musical Genre* (Ph.D. Thesis). University of Ibadan.
- Idamoyibo, A. A. (2011). Esa music and the struggle for relevance in the 21st century. *Pakistan J. Soc. Sci.* 8, 234–239. doi: 10.3923/pjssci.2011.234.239
- Inwald, E. C., Döllinger, M., Schuster, M., Eysholdt, U., and Bohr, C. (2011). Multiparametric analysis of vocal fold vibrations in healthy and disordered voices in high-speed imaging. *J. Voice* 25, 576–590. doi: 10.1016/j.jvoice.2010.04.004
- Iverson, G. K., and Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology* 12, 369–396. doi: 10.1017/S0952675700002566
- Jessen, M., Koster, O., and Gfroerer, S. (2005). Influence of vocal effort on average and variability of fundamental frequency. *Int. J. Speech Lang. Law* 12, 174–213. doi: 10.1558/sll.2005.12.2.174
- Johnstone, T., van Reekum, C. M., Hird, K., Kirsner, K., and Scherer, K. R. (2005). Affective speech elicited with a computer game. *Emotion* 5, 513. doi: 10.1037/1528-3542.5.4.513
- Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* 129, 770. doi: 10.1037/0033-2909.129.5.770
- Kager, R. (2007). “Feet and metrical stress,” in *The Cambridge Handbook of Phonology*, ed P. de Lacy (Cambridge: Cambridge University Press), 195–227.
- Kamiloglu, R. G., Fischer, A. H., and Sauter, D. A. (2020). Good vibrations: a review of vocal expressions of positive emotions. *Psychon. Bull.* 27, 237–265. doi: 10.3758/s13423-019-01701-x
- Kenstowicz, M. (2006). “Tone loans: the adaptation of English loanwords into Yoruba,” in *Selected Proceedings of the 35th Annual Conference on African Linguistics*, eds J. M. Mugane, J. P. Hutchison., and D. A. Worman (Somerville, MA: Cascadilla Proceedings Project), 136–146.
- Khan, S., and u. D. (2012). The phonetics of contrastive phonation in Gujarati. *J. Phon.* 40, 780–795. doi: 10.1016/j.wocn.2012.07.001

- Kim, C.-W. (1970). A theory of aspiration. *Phonetica* 21, 107–116. doi: 10.1159/000259293
- Koenig, L. L., and Fuchs, S. (2019). Vowel formants in normal and loud speech. *J. Speech Lang. Hear. Res.* 62, 1278–1295. doi: 10.1044/2018_JSLHR-S-18-0043
- Koller, W., O'Hara, R., Dorus, W., and Bauer, J. (1985). Tremor in chronic alcoholism. *Neurology* 35, 1660–1660. doi: 10.1212/WNL.35.11.1660
- Ladefoged, P., and Johnson, K. (2015). *A Course in Phonetics*. Wadsworth, OH: Cengage Learning.
- Lenth, R., and Lenth, M. R. (2018). Package 'lsmeans'. *Am Stat.* 34, 216–221.
- Lieberman, M., and Prince, A. (1977). On stress and linguistic rhythm. *Linguist. Inquiry* 8, 249–336.
- Lindquist, K. A., Gendron, M., Satpute, A. B., and Lindquist, K. (2016). "Language and emotion," in *Handbook of Emotions*, eds M. Lewis, J. M. Haviland-Jones, and L. F. Barrett, 4th Edn (New York, NY: The Guilford Press), 579–594.
- Madill, C., Nguyen, D. D., Yick-Ning Cham, K., Novakovic, D., and McCabe, P. (2019). The impact of nasalance on cepstral peak prominence and harmonics-to-noise ratio. *Laryngoscope* 129, E299–E304. doi: 10.1002/lary.27685
- Martins, R. H. G., Benito Pessin, A. B., Nassib, D. J., Branco, A., Rodrigues, S. A., and Matheus, S. M. M. (2015). Aging voice and the laryngeal muscle atrophy. *Laryngoscope* 125, 2518–2521. doi: 10.1002/lary.25398
- McKenna, V. S., and Stepp, C. E. (2018). The relationship between acoustical and perceptual measures of vocal effort. *J. Acoust. Soc. Am.* 144, 1643–1658. doi: 10.1121/1.5055234
- McPherson, L., and Ryan, K. M. (2018). Tone-tune association in Tommo So (Dogon) folk songs. *Language* 94, 119–156. doi: 10.1353/lan.2018.0003
- Nix, J., Perna, N., James, K., and Allen, S. (2016). Vibrato rate and extent in college music majors: a multicenter study. *J. Voice* 30, 756–e31. doi: 10.1016/j.jvoice.2015.09.006
- Ọgúndẹ́jì, P. A. (1991). *Introduction to Yoruba Oral Literature*. Ibadan: Ibadan Center for External Studies; University of Ibadan.
- Okpewho, I. (1992). *African Oral Literature: Backgrounds, Character, and Continuity*, Vol. 710. Bloomington, IN: Indiana University Press.
- Ọlabimtan, A. (1977). Rhythm in Yoruba poetry: the example of orin-arungbe. *Res. Afr. Literat.* 8, 201–218.
- Olajubu, C. O., and Ojo, J. R. O. (1977). Some aspects of Ọyó Yorùbá masquerades. *Africa* 47, 253–275. doi: 10.2307/1158862
- Olajubu, O. (1974). Iwí Egúngún chants: an introduction. *Research in African Literatures*, 31–51.
- Olajubu, O. (1984). Funeral dirges of Yoruba hunters. *Res. Afr. Literat.* 14, 592–596.
- Olaniyan, S. O. (2013). An ecocritical reading of Ijala chant: an example of Ogundare Foyanmu's selected Ijala chant. *J. Literat. Art Stud.* 3, 692–701. doi: 10.17265/2159-5836/2013.11.004
- Ọlátúnjí, D. O. (1979). The Yoruba oral poet and his society. *Res. Afr. Literat.* 10, 179–207.
- Ọlátúnjí, M. O. (2012). Modern trends in the Islamized music of the traditional Yorùbá concept, origin, and development. *Matatu* 40, 447–455. doi: 10.1163/18757421-040001029
- Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., et al. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *Am. J. Speech Lang. Pathol.* 27, 887–905. doi: 10.1044/2018_AJSLP-17-0009
- Phadke, K. V., Laukkanen, A.-M., Ilomäki, I., Kankare, E., Geneid, A., and Švec, J. G. (2020). Cepstral and perceptual investigations in female teachers with functionally healthy voice. *J. Voice* 34, 485–e33. doi: 10.1016/j.jvoice.2018.09.010
- Pulleyblank, D. (1986). *Tone in Lexical Phonology*. Dordrecht: D. Reidel Publishing Company.
- Pulleyblank, D. (1988). Vocalic underspecification in Yoruba. *Linguist. Inquiry* 19, 233–270.
- Pulleyblank, D. (2009). "Yoruba," in *The World's Major Languages*, ed B. Comrie (Oxford: Taylor & Francis), 866–882.
- Purvis, T. M. (2009). Speech rhythm in Akan oral praise poetry. *Text Talk* 29, 201–218. doi: 10.1515/TEXT.2009.009
- Riebold, J. (2013). *Vowel Analyzer*. Available online at: <https://github.com/jmriebold>
- Rosenthal, A. L., Lowell, S. Y., and Colton, R. H. (2014). Aerodynamic and acoustic features of vocal effort. *J. Voice* 28, 144–153. doi: 10.1016/j.jvoice.2013.09.007
- Rumsey, D. J. (2009). *Statistics II for Dummies*. Hoboken, NJ: Wiley Publishing Inc.
- Schellenberg, M. (2009). "Singing in a tone language: Shona," in *Selected Proceedings of the 39th Annual Conference on African Linguistics*, eds A. Ojo and L. Moshi (Somerville, MA: Cascadia Proceedings Project), 137–144.
- Schellenberg, M. H. (2013). *The Realization of Tone in Singing in Cantonese and Mandarin* (Ph.D. Thesis). University of British Columbia.
- Scherer, K. R. (1985). "Vocal affect signaling: a comparative approach," in *Advances in the Study of Behavior*, Vol. 15, eds K. Scherer, J. Rosenblatt, C. Beer, M. Busnel, and P. Slater (New York, NY: Academic Press), 189–244.
- Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychol. Bull.* 99, 143. doi: 10.1037/0033-2909.99.2.143
- Scherer, K. R. (2021). Comment: advances in studying the vocal expression of emotion: current contributions and further options. *Emot. Rev.* 13, 57–59. doi: 10.1177/1754073920949671
- Scherer, K. R., Grandjean, D., Johnstone, T., Klasmeyer, G., and Bänziger, T. (2002). "Acoustic correlates of task load and stress," in *7th International Conference on Spoken Language Processing (Interspeech 2002)*, eds J. H. L. Hansen and B. Pellom (Adelaide, SA: Causal Productions), 2017–2020.
- Scherer, K. R., Sundberg, J., Fantini, B., Trznadel, S., and Eyben, F. (2017). The expression of emotion in the singing voice: acoustic patterns in vocal performance. *J. Acoust. Soc. Am.* 142, 1805–1815. doi: 10.1121/1.5002886
- Seashore, C. E. (1938). *The Psychology of Music*. New York, NY: McGraw-Hill.
- Seyfarth, S., and Garellek, M. (2018). Plosive voicing acoustics and voice quality in Yerevan Armenian. *J. Phon.* 71, 425–450. doi: 10.1016/j.wocn.2018.09.001
- Siertsema, B. (1959). Stress and tone in Yoruba word composition. *Lingua* 8, 385–402. doi: 10.1016/0024-3841(59)90037-3
- Sundberg, J. (1977). The acoustics of the singing voice. *Sci. Am.* 236, 82–91. doi: 10.1038/scientificamerican0377-82
- Sundberg, J., Salomão, G. L., and Scherer, K. R. (2021). Analyzing emotion expression in singing via flow glottograms, long-term-average spectra, and expert listener evaluation. *J. Voice* 35, 52–60. doi: 10.1016/j.jvoice.2019.08.007
- Tolkmitt, F., Helfrich, H., Standke, R., and Scherer, K. R. (1982). Vocal indicators of psychiatric treatment effects in depressives and schizophrenics. *J. Commun. Disord.* 15, 209–222. doi: 10.1016/0021-9924(82)90034-X
- Trautmüller, H., and Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* 107, 3438–3451. doi: 10.1121/1.429414
- Uzochukwu, S. (1981). *Traditional Elegiac Poetry of the Igbo: A Study of the Major Types* (Ph.D. Thesis). University of Lagos.
- Villepastour, A. (2010). *Ancient Text Messages of the Yorùbá Bātá Drum: Cracking the Code*. Surrey: Ashgate Publishing Limited.
- Wasserstein, R. L., and Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *Am. Statist.* 70, 129–133. doi: 10.1080/00031305.2016.1154108
- Wolfe, V., and Martin, D. (1997). Acoustic correlates of dysphonia: type and severity. *J. Commun. Disord.* 30, 403–416. doi: 10.1016/S0021-9924(96)00112-8
- Xu, Y. (2013). "Prosodypro—A Tool for Large-Scale Systematic Prosody Analysis," in *Tools and Resources for the Analysis of Speech Prosody* eds B. Bigi and D. Hirst (Aix-en-Provence: Laboratoire Parole et Langage), 7–10.
- Yai, O. (1989). "Issues in oral poetry: criticism, teaching, and translation," in *Georgetown University Round Table on Languages and Linguistics 1986*, eds Battestini and P. X. Simon (Washington, DC: Georgetown University Press), 91–106.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.