1

Representation of Yorùbá Tones by a Talking Drum
An Acoustic Analysis

Samuel Akinbo
University of British Columbia, Vancouver, Canada

Abstract
The present paper proposes an articulatory and acoustic study of the representation of Yorùbá tones in gángan (a talking drum). The video and spectrographic analyses of the data collected from five native drummers in Nigeria show that the number of syllables in a word directly corresponds to the number of strikes on the drum membrane. As the talking drum resonates from the strikes, the drummers tightened and loosened the drum membrane to articulate the three tones in Yorùbá. Furthermore, tonal processes such as tone contour formation on the second tone in HL or LH sequences are musically rendered. Based on this evidence, this paper concludes that drummers are able to represent syllables, lexical tones and tonal processes of Yorùbá speech with a talking drum.

Keywords
lowering, music, phonetics, raising, syllable, talking drum, tone, Yorùbá.

Résumé
Cet article propose une étude articulatoire et acoustique de la façon dont les tons du yorùbá sont exprimés au moyen du gángan (tama ou tambour parlant). Des analyses vidéo et spectrographiques portant sur les données recueillies auprès de cinq tambourinaires nigérians de langue maternelle yorùbá montrent que le nombre de syllabes d’un mot donné correspond exactement au nombre de fois que la membrane du tambour est frappée. De plus, pendant que le tambour résonne, les tambourinaires tendent ou relâchent alternativement la membrane afin de produire les trois tons existant en yorùbá.
En outre, des phénomènes tonals tels que les modulations apparaissant sur le second ton des séquences HL ou LH sont également exprimés musicalement. Il en découle que ces tambourinaires sont capables de transposer au moyen de leurs tama la structure syllabique du yorùbá, les tons lexicaux de cette langue ainsi que certains des phénomènes tonals qui la caractérisent.

Mots clés
musique, phonétique, syllabe, tama, tambour parlant, ton descendant, ton montant, yorùbá.

1. Introduction
Yorùbá (Kwa, Benue-Congo) has three contrastive tones:¹ H(igh) bà ‘to meet’, L(ow) bà ‘to land’, and M(id) ba ‘to plait’ (Ward 1952; Bamgbose 1966 & 2000; Awobuluyi 1978). In addition to the three contrastive tones, a combination of two of the tones can form, among others, the following tone contours: ríí H M ‘see him/her/it’; rɛɛ L H ‘tired him/her’; jẹẹ M H ‘ate it’.²

Vowel initial words in Yorùbá must begin with L or M (Akinlabi 1985; Pulleyblank 1988). With sequences of HL and LH in Yorùbá, a contour tone is formed on the second tone. However, this does not happen in HM, MH, LM or ML sequences (Ward 1952). Consequently, words like /kʊpákò/ ‘chew stick’ and /ilú/ ‘city’ are realised as [kʊpá] and [ilú] respectively. Furthermore, the pitch value of H is significantly higher in a sequence of HL but not HM (Akinlabi & Liberman 1995). For example, the pitch of H in [bájɔ] ‘Báyò (name)’ is higher than the initial H in [bájɔ] ‘exit through (v.)’.

It is widely known that these level tones and contour tones can be musically rendered in the notes of gángan, an hourglass shaped pressure drum that belongs to a family of drums called dundún (Arewa & Adekola 1980; Euba 1990; Villepastour 2016; Agawu 2016). This attribute of gángan is comparable to those of other Yorùbá drums (e.g. Bátá, Villepastour 2016) and the speech surrogacy system in other cultures (Bagemihl 1988; Agawu 1995 & 2016; McPherson 2019). While most of the research on Yorùbá talking drums espouses socio-cultural and musical perspectives (Beier

1. The funding for this research is from a SSHRC insight grant (435-2016-0369) awarded to Douglas Pulleyblank. For making a recording space available, I thank Demola Lewis and the management of Diamond FM, University of Ibadan. For helpful comments and guidance on various aspects of this work, I thank Douglas Pulleyblank, Rose-Marie Déchaine, Gunnar Hansson, Tolu Odebunmi, Oksana Tkachman, Emily Sadlier-Brown, the anonymous reviewer and the audience at ACAL50 2019. Errors of fact or explanation are my own responsibility.

2. Transcriptions are in standard Yorùbá orthography throughout, unless an example is enclosed in phonemic slants “/ /” or phonetic square brackets “[ ]”. In Yorùbá orthography, [ɛ] = ọ, [ɛ] = ẹ, long vowel = double letters (e.g. e = short and ee = long), nasalised vowel = Vn, [J] = ʂ, [kp] = p, H = , L = , M = unmarked for tone, syllabic nasal = a tone-marked nasal. Tone is indicated throughout, including tonal alternations.
1954; Armstrong 1954; Apter 1985; Euba 1990; Ajayi 1992; Villepastour 2016), the present work is developed within a linguistic perspective.

This work addresses the following questions: (1) How are Yorùbá tones articulated with the drum? (2) Do the drummers represent other linguistic information in addition to tone? (3) What are the acoustic correlates of tone in talking drum notes? (4) Are tonal processes, i.e. raising the pitch of H tone in HL or contour formation on the second tone in HL and LH sequences, also represented?

To address these issues, this paper proposes an articulatory and acoustic study on the representation of Yorùbá tones by a talking drum. For ease of comparison with Yorùbá tones, I refer to the tones of Yorùbá words rendered by a talking drum as drum tones.

2. The talking drum: gángan and its components

Gángan is an hourglass-shaped pressure drum used by the Yorùbá people of southwestern Nigeria. Given that people communicate with the drum, the Westerners referred to it as “talking drum”. The drum’s communicative capability is based on the three tones of Yorùbá. With Yorùbá being the source language of the drum, it is designed to produce notes which are analogous to Yorùbá tones (Euba 1990; Villepastour 2016).

Gángan has six components, namely a wooden barrel, two surface membranes, tension cords, a taut rope, a strap, and a curved stick. The wooden barrel is carved into an hourglass shape with a tunnel-like hole linking the two ends. There are two membranes in this drum. The membranes which are made from the skin of a goat’s foetus are stretched with threads over the holes of the wooden barrel. These membranes covering the two edges of the drum are connected and tightly held together with cords which are leather strings. These cords run along the body of the drum. Given the fact that the cords are connected to the two membranes of the drum, the membranes can be stretched by tensioning the cords. It is required for the membranes of the drum to be loose or slack on the drum barrel before drumming. However, when the drum has not been used for a while, the drum membrane shrinks and becomes tightly held to the barrel. In such a case, to make the drum suitable for music or communication, the taut rope is used in order to expand the membranes by compressing the cords for at least four hours. The taut rope can be made from wool or leather. A curved stick is used for striking the drum. In order to avoid puncturing the drum, the head of the stick is covered with flat rubber. The drum is suspended over the left or right shoulder of the drummer with a strap and the drum hangs lower by the armpit. This strap is made from a padded cloth. These components are shown in Figure 1, adapted from Arewa & Adekola (1980); the labels are based on
terms from Villepastour (2016: 80-81). The image has been modified for an accurate description of the drum. The drum emits varying pitches when one of its membrane is struck and stretched at varying degrees.

3. Methodology

Data were elicited from five male drummers who are native speakers of Yorùbá at Diamond FM radio station, University of Ibadan, Nigeria. The data were recorded in a soundproofed room with a SHURE WH30XLR cardioid condenser (a headset microphone) and a Rode NGT2 supercardioid condenser (a shotgun microphone) at the sampling rate of 48.1 kHz in .wav format. The microphones were attached to a zoomQ8 camcorder. The headset microphone captured the speech of the drummer and the shotgun microphone was pointed at the drum. The audio from the two microphones were saved as separate files at the same time as video files. All the participants in the study used the same drum except for one who insisted on using his own drum.

The stimuli in this study are grouped into level tones and contour tones. The stimuli with level tones have monosyllabic and disyllabic words, which are shown in Table 1. The monosyllabic words with level tones are three words with each word bearing either H, L, or M. The disyllabic words with level tones cover nine tonal combinatorial possibilities.
Table 1 — Tone types in mono- and disyllabic stimuli

<table>
<thead>
<tr>
<th>Syl</th>
<th>H</th>
<th>M</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

The stimuli with contour tones contain three tone-types, namely L̃H, H̃M, and M̃H tones. These contour tones are from three monosyllabic words with two morae each: rë̀é ‘tired it’; rii ‘see it’; jë́́ ‘eat it’.³

In total, the stimuli in all the groups contain 15 words (i.e. 3 monosyllabic words + 9 disyllabic words + 3 contour-toned words). Each of the drummers drummed each stimulus at least thrice. In other words, there are 45 tokens (15 x 3) for each drummer. Although drumming in a speech surrogacy system involves 3 drum modes, namely a speech mode, a signal mode and a dance mode (Agawu 1995: 91), this study is solely based on speech mode drumming.

4. Results
The visual observations were based on the data from all the drummers. For spectrographic analyses, drum tones were segmented manually in Praat (Boersma 2001). A Praat script was used to extract F0 values (Xu 2013). Among all the sets of data elicited, only data from 3 of the 4 participants who used the same drum were instrumentally analysed. This was done in order to control for the acoustic properties of the drum. Data from the fourth participant who used the same drum were not included because, being a young drummer, he could not drum all the words.

³ In Orie & Pulleyblank (2002), it is argued whether some bimoraic syllables, which are derived through a hiatus resolution, contain a syllable or two. For the token used in this study, we assume here that the combination of a CV verb and a 3SG.OBJ pronoun is monosyllabic.
4.1 Articulatory result
To drum a monosyllabic word, the drummers struck the membrane once and twice for a bisyllabic word. To drum a monosyllabic word with two moras, the drummers also struck the membrane once. To drum H tone, the drummers tightly compressed the tension cords under their arm then struck the membrane. For M tone, they loosely compressed the tension cords then struck the membrane. In drumming L tone, the membrane was struck without compressing the tension cords. To drum H̃M contour, the drummers tightly compressed the tension cords then struck the membrane for the initial H of the contour. As the membrane was vibrating from the strike, the tension cords were then slightly loosened for M. In drumming M̃H, the tension cords were slightly compressed then the membrane was struck for the initial M. As the membrane was vibrating from the strike, the tension cords were tightly compressed for the second H of the contour tone. To drum L̃H, the drummer struck the membrane without compressing the tension cords for the initial L. As the membrane was vibrating, the tension cords were tightly compressed for the H of the contour.

4.2 Acoustic result
Spectrograms show pitch is the main property that distinguishes a drum tone, so the analyses solely focus on pitch. The pitch tracks of level drum tones (H, L and M) and contour drum tones (H̃M, M̃H, and L̃H) in isolation are shown in Figures 2 and 3 respectively. The pitch tracks figures are from the same drummer. The pitch tracks figures clearly show a direct correlation between H, L, and M tones and their different pitches.

Since each drum tone was drummed the same way by the 3 participants, I combined all the pitches of each drum tone. The mean values of each drum tone are shown in Table 2. To mitigate the impact of large outliers and aggravate the impact of small ones, the mean values is from a harmonic mean. As shown in Table 2, H has a higher F0 than M, and M a higher F0 than L.

<table>
<thead>
<tr>
<th>Drum tone</th>
<th>F0 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>172.307</td>
</tr>
<tr>
<td>M</td>
<td>124.7897</td>
</tr>
<tr>
<td>L</td>
<td>61.42983</td>
</tr>
</tbody>
</table>
4.3 Acoustic result for sequences of tone

With sequences of HL and LH drum tones shown in Figures 4 and 5 respectively, a tonal contour is formed on the second drum tone. For instance, a high-falling contour is formed on the second tone in the drumming of pákò 'chewing stick', while a low-rising is formed on the second tone in the drumming of ́ilú 'city'. However, this does not happen in LM and ML shown in Figures 6 and 7 respectively.
To check if there is any effect of an adjacent drum tone, the F0 values of each tone in a sequence of two tones were boxplotted in Figure 8 by using ggplot2 (Wickham 2016). Given the contour tone formation on the second tone in HL and LH sequences, only F0 values of word-initial H, L and M drum tones preceding H, L and M were boxplotted. These F0 values were extracted at the 50% point.4

Figure 8 shows a clear distinction between H L and M drum tones. For the H drum tone, the F0 value is higher when it precedes an L drum tone. The F0 value of an L drum tone is lower when it precedes an H drum tone. As for M drum tone, it has a higher F0 value when it precedes an M drum tone.

In R (R Core Team 2018),5 the linear mixed effects (LME)6 model (Bates et al. 2015) was fitted to F0 values of the word-initial drum tones to determine the significant effect of the following tones. For LME, the dependent variable was F0 values taken at the 50% point; the fixed effects were the following tones; the variable subjects (i.e. the individual realization of each drummer) was included as the random effect. For fixed effect factor, a following tone which is like the preceding tone was included as the reference level for each tone, i.e. H for H, L for L and M for M.

4. The y-axis contains the F0 of the drum tone while the distributional characteristics of the F0 range of each tone in an environment is represented in the box. The box contains the middle 50% of the F0 range. The mid-line in the box marks the median. The top cell of the box contains the maximum 25% of the box and the bottom cell of the box contains the minimum 25% of the box. Each of the lines below and above the box contains 25% of the F0 values which are outside the middle 50%. The dots represent the outliers in the data.

5. R is a free software for statistical analysis and graphics (URL: https://www.r-project.org).

6. LME describes the relationship between a response variable and independent variables, with coefficients that can vary with respect to one or more grouping variables.
The result shows that the following drum tone had a significant effect on the pitch of the preceding drum tone. For example, the F0 of H tone was significantly higher when it precedes an L tone ($t = 3.3$, $p = 0.0029$). On the other hand, the following M ($t = 0.5$, $p = 0.6$), relative to the following H, had no significant effect on the F0 of the preceding H tone.

Furthermore, the F0 of L tone was significantly lower when it precedes an H tone ($t = 10$, $p = 0.0008$). However, the following L ($t = 0.05$, $p = 0.96$), relative to the following H, had no significant effect on the F0 of the preceding L tone.

Finally, the F0 of M tone was significantly higher when it precedes an M tone ($t = 2.1$, $p = 0.04$). However, the following L ($t = -0.04$, $p = 1$), relative to the following M, had no significant effect on the F0 of the preceding M tone.

---

7. $t$: test statistic; $p$: probability value.
5. Summary, discussion and conclusion

This study shows that the number of strikes the drummers make on the drum membrane directly corresponds to the number of syllables in the presented stimulus. As the drum vibrates from the strikes, the drummer varied the compression of the tension cords in articulating the H, M and L tones in Yorùbá. The acoustics result shows H, M and L drum tones had distinctive pitches. In sequences of HL and LH drum tones, a contour was formed on the second drum tone. The pitch of the H drum tone is significantly raised in a sequence of HL drum tones. However, the pitch of L drum tone is significantly lower in a sequence of LH drum tones.

The pitch distinction between H, L, and M drum tones support previous findings that Yorùbá tones are distinctively rendered with talking drum (Euba 1990; Villepastour 2016). The direct correspondence between the number of syllables in a word and the striking of the drum membrane also suggests that the drummers render the number of syllables in a word with the talking drum. The realisation of a contour tone on the second drum tone in a sequence of HL and LH is similar to the pattern found in Yorùbá speech (Pulleyblank 1988; Akinlabi & Liberman 1995). Raising the pitch of H drum tone in a sequence of HL and lowering the pitch of L drum tone in a sequence of LH in drumming is also similar to the phonetic realisation of tones in spoken Yorùbá (Connell & Ladd 1990; Akinlabi & Liberman 1995). Phonologically, a contrast maximisation for tonal features is considered the functional motivation for the H raising and L lowering in a sequence of HL and LH respectively (Diehl & Lindblom 2004; Yoshida 2011). By extension, it is possibly the case that the H raising and L lowering in the drum tone is a result of the same contrast maximisation. The similarity between drumming and speech tones indicates that the drummers are sensitive to the tonal processes or phonetic properties of speech tones in drumming. An argument for this sensitivity is the conceptualisation of instrumental pairs in some African drumming traditions. For instance, a drum with a low pitch is conceptualised as a male but a drum with a high pitch is conceptualised as a female (Agawu 2016: 83). This conceptualisation can be phonetically corroborated with the results of several studies showing that the pitch range of a female is higher than that of a male (Hollen & Ship 1972; Pegoraro-Krook 1988; Traunmüller & Eriksson 1995).

The contour tone formation on the second syllable in the HL and LH sequences raises an issue on the articulation of tone in drumming considering the fact that the second tone was articulated on the drum without compressing the tension cord: the formation of contours (high falling for HL and low-rising for LH) at the end of the second tone in these sequences (see Figures 4 and 5) should have been unattested since no compression
was made in articulating this second tone. So, (i) the high-falling contour formation at the end of the second drum tone in HL sequence and (ii) the low-rising contour formation at the end of the second drum tone in a LH sequence both suggest that, in addition to varying cord compression, a secondary articulatory strategy is possibly involved in the production of drum tones. Future research on the drumming of Yorùbá words with gängan should focus on this additional strategy.8

Given the fact that this study was based on lexical tones and that tones perform grammatical functions in certain constructions in Yorùbá (Akinlabi 1985; Liberman & Akinlabi 2000; Laniran & Clements 2003), future research should investigate the representation of the grammatical tones in Yorùbá by a talking drum. The result of a language acquisition study in Orie (2006) shows that rhythmic language and leisurely activities aid L2 acquisition of tones. Considering the fact that a drum language is both rhythmic and leisurely (Agawu 1995 & 2016; Orie 2006), research on L2 acquisition should investigate the effect of the perception and production of drum tones on the acquisition of tones.

Be that as it may, this study clearly shows that Yorùbá drummers faithfully render the number of syllables, the lexical tones and the phonetic realisation of Yorùbá tones with a talking drum.

References


AKINLABI, Akinbiyi, 1985, Tonal Underspecification and Yoruba Tone, PhD dissertation, Ibadan, University of Ibadan.


8. This is based on the comments made by Kofi Agawu, Laura McPherson and Murray Schellenberg at ACAL50.


