Language Sciences 98 (2023) 101553

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

Sound-meaning mapping: Verbal imitation of Super Mario music by Yorùbá gamers☆

Samuel Kayode Akinbo

Institute of Linguistics, University of Minnesota, United States

A R T I C L E I N F O

Article history:
Received 3 June 2022
Received in revised form 26 February 2023
Accepted 24 April 2023
Available online xxx

Keywords:
Music
Meaning
Tone
Videogame
Imitation
Social expectation
Context
Iconicity

A B S T R A C T

An aspect of gaming culture among Yorùbá millennials is verbally interpreting certain musical motifs of the popular videogame called Super Mario Bros. The themes of the verbal interpretations are comparable to those of music texts at traditional Yorùbá competitions. Drawing on the Yorùbá music tradition, the account in this work is that, to the gamers, the background music of the videogame performs a similar function as the music at traditional Yorùbá competitions. Semantically, the choice of words in the linguistic interpretation is conditioned by the situational contexts or scenes where the music is heard in the videogame. The results of an acoustic analysis show that the pitch contours of the linguistic interpretations resemble the pitch trajectories of the corresponding music motifs. Thus, the sequence of words in each linguistic interpretation is determined by vocal imitation. This study suggests that the linguistic processing of music does not only involve phonetic iconicity but includes contextual inference and social expectation. The interpretive moves clearly point to strong parallels between sound-meaning mapping in spoken language and music.

© 2023 Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

This paper explores how Yorùbá gamers map lexical meaning to videogame music in their native language. Sound-meaning mapping is an obligatory aspect of spoken language (Hockett, 1960; de Saussure, 1974; Dingemanse et al., 2015) but an optional aspect of music. For example, it has been repeatedly shown that humans associate affective and non-affective meaning to various pitch and tempo patterns (Feld, 1984; Hacohen and Wagner, 1997; Koelsch et al., 2004; Patel, 2008; García et al., 2014; Neumeyer, 2015). Prominent examples of sound-meaning mapping outside spoken language are musical speech surrogates, such as talking drums and flutes of West Africa, Asia and America (Stern, 1957; Carrington, 1971; Ekwueme, 1974; Bradley, 1979; Lo-Bamijoko, 1987; Poss, 2005; Winter, 2014; Poss, 2012; Seifart et al., 2018; McPherson, 2018). In this case, musicians map linguistic meaning to music melodies and rhythm by imitating linguistic features with musical instruments. Consequently, musicians from these cultures are able to communicate in a language-like form via musical instruments.

Studies of music communication in cultures with musical speech surrogates have tended to focus more on musical imitation of language, but it is also important to examine the "opposite", which involves verbal imitation or interpretation of

☆ This research was carried out when I was a student working as a research assistant 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, Gunnar Hansson, Khia Johnson, Timo Roettger, and the reviewers. Errors of fact or explanation are my own responsibility.

E-mail address: samuel.akinbo@ubc.ca.

https://doi.org/10.1016/j.langsci.2023.101553
0388-0001/© 2023 Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

music melodies. Using linguistic instrumentation and methodology, the present study contributes to research on verbal interpretation of musical sounds in a culture with speech surrogates. The present work is based on conventionalised verbal imitations of the music in the Super Mario game by Yorùbá gamers.

I will argue that Yorùbá gamers mapped lexical meaning to the videogame music by (i) verbally imitating the pitch contours of the videogame music motifs, (ii) using the scenes of musical performance in the videogame and (iii) projecting the social expectation about music in Yorùbá culture to the function of the videogame music. This work is of interest for three main reasons. First, the role of vocal imitation in mapping lexical meaning to music melodies suggests a parallel between language and music, considering that sound symbolism in human language (e.g., onomatopoeia) also involves resemblance-based mapping between language and language-external sources (Ramachandran and Hubbard, 2001; Dingemanse et al., 2015; Akinbo, 2021a, 2021c). Most of the evidence for verbal imitation in musical meaning is intra-cultural (Patel and Iversen, 2003; Villepastour, 2014; James, 2021), but the verbal interpretation of the videogame music by Yorùbá gamers presents inter-cultural evidence for verbal imitation in musical meaning. The second area of interest is the role of context in mapping linguistic meaning to language-external sounds. Just as context contributes to meaning in natural language (Eberhard et al., 1995; Tanenhaus and Trueswell, 1995), there is evidence to suggest that context also plays a crucial role in mapping meaning to music (Villepastour, 2014). However, the simulation of context in experimental conditions poses a serious challenge for the study of musical cognition (Valsiner, 1994; Parrott and Hertel, 1999). The present study presents a real-life evidence for the role of context in musical meaning, considering that the scenes of musical performance in the videogame also determines the words that were mapped to the musical melody. Third, the results of this work are consistent with the claim that videogame music supports the perception of game world, player's involvement and game narrative (Zehnder and Lipscomb, 2006; Grimshaw, 2008; Nacke et al., 2010; Sanders and Cairns, 2010). While research on videogame music mostly emphasises music structure and gameworld contexts (Munday, 2007; Laroche, 2012), the present study indicates that cultural background of gamers is also relevant. Most importantly, the overlap between language and music is crucial to the proposal in linguistic theory and cognitive science that language, music and other processes share cognitive properties, such as paying attention to similarities and categories (see Patel, 2012 on music cognition and Archangeli and Pulleyblank, 2022 on the emergence of grammar).

As a background to the discussion on the verbal interpretation, the discussion in §2 focuses on aspects of language and music tradition that form basis of the verbal interpretation. For the present study, I document and analyse the conventionalised linguistic interpretations of some videogame musical motifs by Yorùbá gamers. The linguistic interpretations and the corresponding musical motifs are presented in §3. The discussion in §4 explores why the gamers interpreted the musical motifs into Yorùbá. To compare the pitch contours of the linguistic interpretations to the pitch trajectories of the corresponding musical motifs, native speakers of Yorùbá were recruited for a production experiment. The conventionalised interpretations were presented to the participants with and without the original videogame music. They reproduced the conventionalised arrangements in speech and sung modes. The results of their spoken and sung productions were acoustically compared to the pitch contours of the videogame music motifs. The details of the methodology and results of the acoustic investigation are presented in §5. In §6, a formal analysis of the results are presented. The summary, discussion and the conclusion are presented in §7.

2. Language background: Yorùbá and its talking drum

Yorùbá is a Volta-Niger language, according to a recent classification (Blench, 2019), with more than 20 million speakers in West Africa and most prominently South Western Nigeria. Yorùbá is a tone language, which means pitch contrast brings about lexical or grammatical meaning distinction (Yip, 2002; Hyman, 2018). As shown in (1), the language contrasts three tones, namely H(igh) L(ow) and M(id) (Bamgboṣe, 1967; Awóbùlúyì, 1978). The tone-bearing unit in Yorùbá is a mora, which is projected by a vowel or a syllabic nasal (Pulleyblank, 2004).

(1) Tonal minimal set in Yorùbá

a.	H	rá	'disappear'	ńlọ	'it's going'
b.	L	rà	'buy'	ńjẹ́	'is it?'
c.	M	ra	'rub'	kí nlọ	'I should go'

By transposing speech tones to tunes and syllables to drum strokes, Yorùbá musicians are able to communicate via native and non-native music instruments (Euba, 1990; Waterman, 2000; Villepastour, 2010, 2014; Durojaye et al., 2021b; González and Oludare, 2022). Using acoustic analysis, Akinbo (2019, 2021b) shows that there is a positive correlation between the pitch contours of the speech tones and those of the corresponding musical rendition, as shown in Fig. 1.

Fig. 1 shows the drummed and spoken forms of the Yorùbá phrase in (2). In the graph, the y-axis contains the acoustic measurement of pitch contour in F0(Hz), and the x-axis contains the tones of the Yorùbá phrase. The dark line is for the drum pitch and the grey line is for the speech pitch. As shown in the graph, the pitch contours of the Yorùbá phrase and the corresponding drum rendition have similar trajectories.

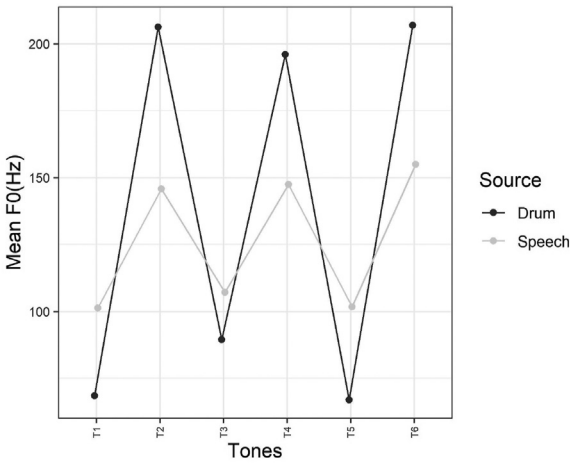

Fig. 1. Mean F0 contours for Yorùbá drum and speech tones (Akinbo 2021, p. 7).

(2) Ògúnbòdé gbà 'gbá
O. take garden egg
"Ògúnbòdé collected a/the garden egg" (Laniran and Clements, 2003)

Nketia (1963) identifies three modes of speech surrogacy, namely speech mode, signal mode and dance mode. These modes are comparable to the three forms of drumming in Euba (1990), which are direct speech form, musical speech mode and song mode. The speech mode "involves direct reproduction of the pitches and rhythm of spoken language" (Agawu 2016:128). Therefore, musicians are capable of musically rendering any verbal utterance in speech mode. The signal mode is similar to the speech mode in terms of matching the pitches and rhythm of spoken language, but the drum messages in this mode are mostly restricted to poetic phrase or epitaph. Unlike the speech and signal modes, the rhythm of the dance mode is musical, "metrically constrained, often affiliated with movement, but not necessarily of linguistic origins" (Agawu 2016:128). However, the rhythms of Yorùbá and its musical speech surrogates have not been subjected to instrumental analysis. In the present study, I only focus on the pitch contours for the verbal imitation of music in Yorùbá.

Regardless of the mode of speech surrogacy, listeners have to decode the linguistic information communicated via the instrument. The verbal interpretation of speech-surrogate messages has not been previously studied but is often alluded to. Previous studies suggest that consumers of speech surrogates rely on pitch contour and context in their interpretation of the messages encoded with speech surrogates (Villepastour, 2010; Sotunsa, 2021), but the possibility of simply interpreting speech-surrogate messages based on previous experience and associated meaning cannot be ruled out. For example, the findings of Durojaye et al. (2021a) indicate that familiarity with speech surrogacy plays a role in distinguishing speech mode from song mode. Though Durojaye et al. (2021a) is about the categorisation of Yorùbá drum stimuli as music-like or speech-like, their finding is relevant to the interpretation of speech-surrogate messages. In fact, the role of familiarity or previous experience in the interpretation of speech-surrogate codes is captured by the Yorùbá maxim which says, òrò àsotì ló ń jẹ́ omo mí gbénà "*if your child understands your code language, it is because you both share the secret*" (Isola, 1982:44).

To control for the effect of familiarity, the present study is based on the verbal interpration of foreign instrumental music which do not involve speech surrogacy. In the next section, I describe the musical source and the verbal interpretations.

3. Verbalising the musical motifs of Super Mario Bros in Yorùbá

The data source for the present study is the conventionalised linguistic interpretation of Super Mario music. In this section, I present background information about the game, the music themes of the game and the conventionalised verbal interpretations of certain musical motifs by Yorùbá gamers.

Nintendo, a Japanese multinational company, released the console videogame called Super Mario Bros in 1985. More than 40 million copies of the game were sold worldwide (Stuart, 2010). The protagonist of the "the jump-and-run" game is Mario who is an Italian plumber. In the game which is set in the Mushroom Kingdom, the player takes on the role of Mario or his brother Luigi in a multiplayer mode. The objective of the game is for Mario (or Luigi) to save Princess Toadstool, but for Mario to save the princess, he needs to survive the main antagonist Bowser, the forces of Bowser, the dangerous terrains in the game and the allotted game time (Nintendo, 1985; Loguidice and Barton, 2012).

The pianist, Koji Kondo, composed the music themes of the game. According to Laroche (2012), the composition of the music themes is a musical reaction to "players' intended impressions". Following Collins (2009: 6), the sounds in Super Mario Bros can be classified into *interactive* and *adaptive* audios. Interactive audios are "sound events directly triggered by the player's input device". The sounds which accompany jumping, hitting an object, shooting, etc. are examples of interactive audios (see Lerner, 2014). On the other hand, adaptive audios are not controlled by the player but are cued by the game scenes, locations, game time, the presence of non-player characters, etc. There are at least eighteen sounds in the game, but the focus of the present paper is four adaptive audio events, namely Overworld/main theme, Flagpole Fanfare theme, Underworld theme and Death theme. Because the Overworld theme is played throughout the game with some tempo variations in certain stages, it is also called the main music theme (see Lerner, 2014; Schartmann, 2015). The musical experience in videogames are the major innovations of Super Mario Bros (Collins, 2009; Lerner, 2014).

The game was very popular in Nigeria around 1990s and early 2000s, but it is uncertain when the game got to the country. An aspect of gaming culture among Yorùbá gamers in Nigeria involves verbally interpreting certain motifs of the videogame music into Yorùbá (Ayoola, 2019). The verbal interpretations developed naturally in the gaming community. Among all the instrumentals in the game, only the adaptive sounds were conventionalised by the gamers. This paper only focuses on the conventionalised interpretations. To my knowledge, the vocal interpretations of Super Mario music by Yorùbá gamers have not been studied until now. The available sources on the interpretations are online posts, which only document fragments of the phenomenon (see the tweets of Akintola 2011 and Odesanya 2013). The present work is based on my documentation of the verbal interpretations, as practiced in South-West Nigeria. Nintendo did not release the music scores of any of their videogame games, including Super Mario Bros. However, authentic scores, such as the one consulted for this work, exist.[1] The music score excerpts containing the interpreted motifs are presented in this section. Each syllable of the verbal interpretations is aligned with the corresponding note of the interpreted motifs.

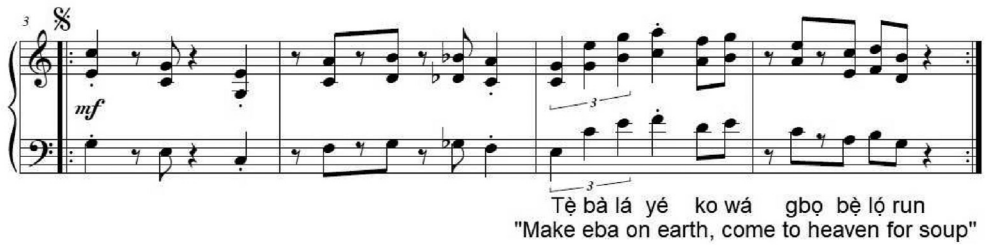

Fig. 2. The interpretation of Overworld motif by Yorùbá gamers.

The interpreted motif of the Overworld theme and its interpretation are presented in Fig. 2. The interpreted motif plays at the very beginning of the first stage (i.e., World 1 level 1). The first instance of the excerpt plays from 389 to 383 game-time countdown, but its repetition plays at other game times. The linguistic interpretation of the relevant Overworld motif is an adaption of a Yorùbá saying which is used as a death threat. The only difference is that the phrase /gún 'yán/ 'pound iyán' in the saying is replaced with the phrase /tè 'bà/ 'make èbà (lit: mash èbà') in the interpretation of the Overworld motif.[2] With the exception of the motif from the Overworld theme, the interpretations of the other musical motifs are original, not based on previously existing Yorùbá sayings.

When Mario drops from the flag after winning the first stage of the game and before entering the castle, the Flagpole fanfare is played. The entire Flagpole fanfare consists three iterations of a motif and a closing chord gesture. Each iterations of the motif has the same verbal interpretation, so only one of the iteration is presented in Fig. 3. While some gamers interpret the closing gestures, others did not. As a result of this, the closing gestures are not included in this work.

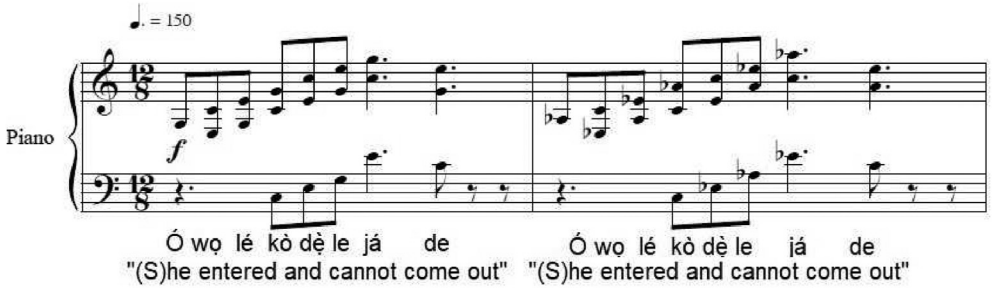

Fig. 3. The interpretation of Flagpole fanfare motif by Yorùbá gamers.

[1] The full score of the excerpts used in this work can be found on https://www.ninsheetmusic.org/browse/series/SuperMario.
[2] Iyán is the Yorùbá name for pounded yam; èbà is a staple food made from dried grated cassava (manioc).

The second stage of Super Mario Bros is set in the underworld, specifically World 1 level 2. Immediately Mario gets into the underworld, the Underworld theme is played. The gamers only interpreted a motif of the Underworld theme, which plays from 400 to 384 game time for its first iteration. The Underworld theme contains four consecutive iterations of the same motif. Fig. 4 presents the first two of the four iterations and their verbal interpretations in Yorùbá. The first and second iterations have separate verbal interpretations. The third and fourth iterations have the same verbal interpretations as the first and second iterations respectively. Other repetitions of the motif are heard throughout the game level and each repetition has the same verbal interpretations as the initial occurrence.

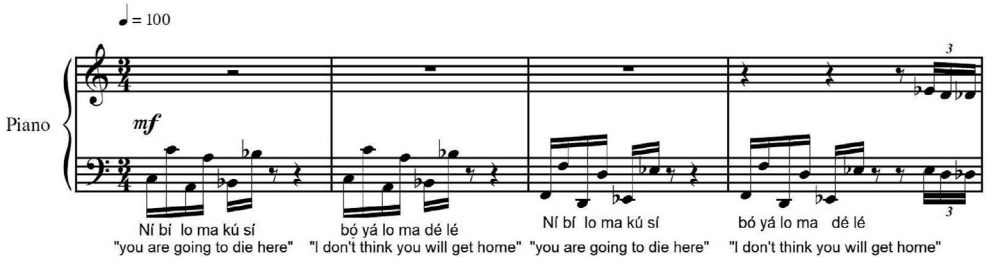

Fig. 4. The interpretation of Underworld motif by Yorùbá gamers.

If Mario dies at any stage of the game, regardless of the cause, the Death theme is played. The gamers also linguistically interpret the motif of the Death theme, which is presented in Fig. 5.

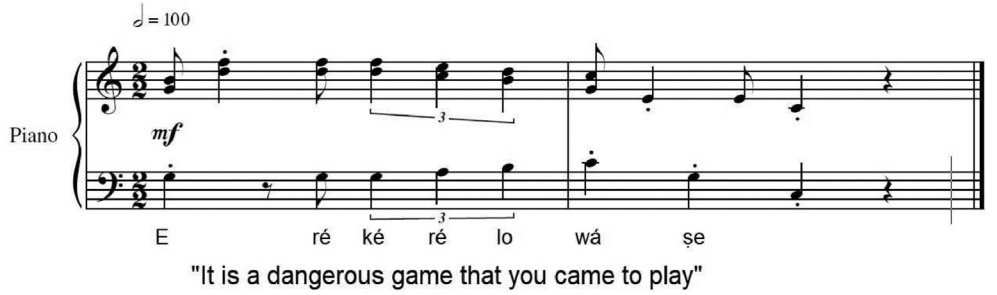

Fig. 5. The interpretation of the Death motif by Yorùbá gamers.

In general, Yorùbá listeners (i.e., Super Mario gamers) are able to linguistically interpret certain motifs of the music themes in Super Mario Bros. The interpretations are usually performed as song. Regardless of whether the interpretation is in speech or song mide, a unifying feature is that a lexical syllable is mapped to each note of the interpreted music motifs. Most Yorùbá gamers who have played Super Mario Bros know these conventionalised interpretations, but it is uncertain whether someone who has not heard the musical motif before might have the same interpretations.

The verbal interpretations of the musical motifs are similar to Yorùbá speech surrogates, considering that a lexical syllable is mapped to each music note of the musical motifs. Studies of Yorùbá speech surrogates show that the pitch contours of the lexical tones with their phonetic realisations are encoded with Yorùbá drums (Villepastour, 2010; Akinbo, 2019). Thus, it is important to know whether the verbal interpretation of music also involves correspondence between verbal and musical pitch contours. In §5, I discuss how the musical contours are mapped to the tone on each syllable of the verbal interpretations. Before focusing on the tone of the syllables, the next section addresses why the gamers considered the background music of the game to be communicative.

4. Strategies for mapping meaning to music

The discussion so far has shown that the gamers linguistically interpreted certain musical motifs of Super Mario by mapping lexical syllables to the music notes of the motifs. Instead of lexical syllables, the Yorùbá gamers could have mapped non-lexical vocables or nonsense syllables to the notes of the musical motifs, as is done in other cultures (see Hughes, 2000; Mullins, 2014; Weir, 2015). So, why did the gamers use lexical syllables instead of vocables? What determines the segmental and tonal properties of each syllable?

The account in this work is that the gamers utilised the situational contexts in the Super Mario game to determine the segments or words which are mapped to the musical motifs. For example, the phrase "(s)he/it entered and cannot definitely come out" is mapped to the relevant motif of the Flagpole fanfare theme which plays immediately Mario is about to enter the castle, as shown in (3a). If we assume that the linguistic interpretation of the Flagpole motif involves mapping the tone sequence H M H L L M H M to the music notes of the Flagpole motif, any phrase with the same tone sequence, such as the examples in (3b-d), could have been mapped to the music motif. Mapping the phrase "(s)he/it entered and cannot definitely come out" to the Flagpole motif suggests that the linguistic interpretation of the musical motif is conditioned by the context of Mario "entering the castle and not coming out".

(3) Possible interpretations of the Flagpole motif
Tone: H M H L L M H M

a.	**Ó wo'lé kòdè le jáde**	'he/she/it entered, cannot come out'
b.	Ó ta mí tètè lo w'ágbo	'he/she/it stung me, quickly look for herbs'
c.	Déwolé àgbàdo dára	' Déwolé, corn is good'
d	etc.	

Similar to the interpretation of the Flagpole motif, the situational context in the game is also utilised in the interpretation of the relevant motif of the Death theme. For instance, the Death theme plays when Mario dies. Mapping the phrase "it is a dangerous game that you came to play" to the motif of the Death theme correlates with mocking the death of Mario for playing a dangerous game, as shown in (4a). The theme of danger is also echoed in the linguistic interpretations of the Overworld motif, as shown in (4b), and the Underworld motif, as shown in (4c).

(4) The interpretations of other Super Mario motifs

a. Death motif	
Tone: M H H H M H M	
erékéré lo wa ṣe	'it is a dangerous game that you came to play'
b. Overworld motif	
Tone: H M H L L M H M	
tèbà láyé ko wá gbobè l'órun	'make èbà on earth, come to heaven for soup'
c. Underworld motif	
Tone: H H M M H H	
níbí lo ma kú sí	'you are going to die here…'
bóyá lo ma dé 'lé	'(I) don't think you will get home'

As mentioned earlier, the overworld theme is the main music theme of the game, and it is the music theme that is heard at the very beginning of the game. The interpretation of the Overworld motif, which is "make èbà on earth and come to heaven for soup" (4b), is an adaptation of a Yorùbá idiom that is used as a death threat. Considering that the objective of the game is for Mario to save the Princess or die trying, the gamers possibly interpreted the Overworld motif as a warning or threat in order to create tension for the player-controlled character (i.e. Mario or Luigi) or more specifically the player. If Mario survives the first level of the game, the Underworld motif plays immediately at the beginning of the second level which is set in the Underworld. Interpreting the first iteration of the Underworld motif as "you are going to die here" is also a threat to the player who escaped death at the previous stage of the game.

The assumption that the interpretations involve mapping a specific tone sequence to the motifs can be supported if we consider that the two iterations of the Underworld motif have the same sequences of tones but different words in their interpretations, as shown in (4c). While the interpretation of the first iteration threatens that Mario is going to "die" at the Underworld level of the game, the interpretation of the second iteration casts doubts on the possibility of Mario repeating the previous feat in the game, which is getting home (i.e., the castle). In this case, getting home is a metaphor for winning. Generally speaking, the interpretation of the first iteration refers to a possible future event in the game and the interpretation of the second iteration connects a previous event (i.e., getting to the castle) with a possible future event. The fact that both interpretations have the same sequences of tones but different words indicates that the gamers are committed to a specific tone sequence in their interpretation and that the same tone sequence can be mapped to two or more phrases, inasmuch as the meaning of each phrase matches the context of the music performance. The tone-sequence requirement is plausibly the motivation for replacing the MH word /iyán/ with the LL word /èbà/ in adapting the popular saying for the verbal interpretation of the Overworld motif (see §3).

As mentioned at the beginning of this section, the gamers could have mapped vocables instead of lexical syllables to the musical motifs. Did the Yorùbá gamers perceive the musical motifs as actual words or phrases? My account is that the Yorùbá gamers perceive the presence of videogame music as the voice of supporters or opponents. For the gamers, the themes of the Overworld and Underworld interpretations are death threats to the player-controlled character or more specifically the player. While the Death motif mocks the death of the player-controlled character, the Flagpole motif is a celebratory motif for the victory of the player. This is interesting, because, with a Western ear, one would rather interpret the music to express Mario's mood (e.g., Overworld motif sounds happy, bright and optimistic) or the atmosphere of the game scenes (as in Whalen, 2004; Laroche, 2012; Schartmann, 2015). But clearly, the Yorùbá gamers hear a different "voice" in the music. This could be described using the theory of musical persona, i.e., a way to understand music by inferring there is a person (or group of persons) speaking through the music (Auslander, 2006; Cochrane, 2010; Fairchild and Marshall, 2019).

The fact that the gamers perceive a voice in the videogame motifs can be considered an effect of their cultural background, given that the interpretation of the Overworld motif is an adaptation of a traditional Yorùbá proverb. Another argument for the effect of cultural background is that the linguistic interpretations of the videogame motifs are thematically comparable to the background music of traditional Yorùbá games and entertainment. Like the interpretation of the videogame motifs, music in Yorùbá and other African cultures is functional, contains textual components, and plays an important role in traditional games and events (Adedeji, 1972; Apter, 1998; Agawu, 2001; Green, 2005; Omojola, 2011; Campbell, 2015; Agawu, 2016). For instance, in Yorùbá societies, instrumental and vocal music often feature in wrestling matches such as those of Oròyèyé festival in Ayédé, Èkìtì State, Nigeria. The background music is played by the supporters from the opposing sides. The music either (5a) threatens, (5b) warns or (5c) mocks the wrestlers. The general goal of the music is to deter the opponent from being victorious. As Apter (1998, p. 77) rightly puts it, [the wrestlers] fought with their bodies, [but their supporters] fought with words. Similar behaviours are found in sport banters of non-African events and games (see Lee, 1985; McLeod, 2006; Vale and Fernandes, 2018).

5) Songs from Oròyèyé festival Wrestling matches (source Apter, 1998)

a	Threat	bírí gbe.	'be carried by the wind'
		òyì gbe …	'be dizzy'
b.	Warning	omodé é gun'gi ògèdè,	'a child that climbs the banana tree,'
		a yó báárá, ase wì …	'will slide down'
b.	Mockery	Adé ró di,	'Adé has thumped down,'
		bí oní koyín	'like a bunch of palm kernels'

Instruments such as the "talking drum" also play similar roles as vocal music at traditional Yorùbá events. For example, during masquerade festivals, Yorùbá drummers often play provocative phrases on the talking drum in order to excite the masquerades or the followers. When the phrases in (6) are played on the drum (or spoken by the spectators), the followers or the masquerades intensify their dance or any other action. In this sense, vocal and instrumental music at traditional Yorùbá events are discourse (Agawu, 2001; Villepastour, 2010, 2014; Agawu, 2016).

(6) Yorùbá drum phrases (Fámúle, 2018)

a	o ò le se bí baba re ń se	'you are not as competent as your father'
b.	b'óbá se pé'mi nì'wo ni	'If I were you'
	ń bá fápá jó, fápá jó …	'I will dance with my hands unceasingly'

Based on the thematic similarities between the verbal interpretations of the musical motifs and the musical text at traditional Yorùbá events, we can say that the gamers consumed the videogame music using their background knowledge about the functions of background music in Yorùbá recreational activities with a competitive component. In other words, the interpretive move of the gamers can be considered an effect of social expectation, which is "an internalised social norm for individuals and organisations…about what people should do" (Hasegawa et al., 2007, p. 180). That the Yorùbá gamers utilised their cultural knowledge about game music in their interpretation of the videogame music is in line with the enculturation account of musical interpretation, which suggests that musical meaning is determined by cultural convention and social background (Keil and Feld, 1994; Walker, 1996; Gregory and Varney, 1996). For example, Feld (1984) suggests that a listener might relate a musical object or event to personal and social conditions, and related experiences where a similar sound object can be heard. Gregory and Varney (1996) argue that "the interpretation of music is determined more by cultural tradition than by the inherent qualities of the music". Social expectations have also been shown to facilitate comprehension and evaluation of spoken language (Rubin, 1992; Devos and Banaji, 2005; Kang and Rubin, 2009; Yi et al., 2013; Babel and Russell, 2015; McGowan, 2015).

Another factor that could be at play is that music in videogames is comparable to leitmotif or film music, i.e., they both involve associating musical phrases to objects, events, people and storytelling in an audio-visual medium (Wagner, 1964; Whalen, 2004; Munday, 2007), and that audiences from diverse cultures have learned to understand film music as being related to the scene, events, and emotions shown in the movie (Thackway, 2002; Cohen, 2011). Notably, Koji Kondo and other composers of videogame music liken the function of the videogame music to leitmotifs in films (Laroche 2012:29). For the interpretation of the Super Mario motifs, it is plausible that the Yorùbá gamers also extended their understanding of film music to videogame music.

In addition to contextual clues and social expectation, the physical properties of music such as pitch and tempo have also been shown to prime the interpretation of music. For example, studies indicate that faster tempo tends to be associated with joy and happiness for Western adult listeners (Scherer and Oshinsky, 1977; Kellaris and Kent, 1993; Gagnon and Peretz, 2003; Webster and Weir, 2005; Eerola et al., 2013). Similar results are found for children music listeners (Dalla Bella et al., 2001; Mote, 2011). However, the studies are based on music traditions without speech surrogates. Considering that Yorùbá has a music tradition with a speech surrogate system that mostly relies on imitating the pitch contours of speech tones (Villepastour, 2010, 2014; Akinbo, 2019), it is important to inquire whether the pitch contours of music motifs also influence the corresponding linguistics interpretation. For this purpose, we must compare the pitch contours of the Super Mario motifs to those of their conventionalised linguistic interpretation in Yorùbá. This investigation is crucial to the tone sequences of the

phrases that are mapped to the videogame motifs. In order to compare the pitch contours of the musical motifs and those of the corresponding linguistic interpretations, I conducted an acoustic investigation.

Before turning to the acoustic investigation, the summary is that, to the Yorùbá gamers, the videogame music is communicative as a result of drawing a parallel between the roles of background music in traditional Yorùbá games and the videogame. In this case, the cultural background and experience of the gamers motivate the linguistic interpretation of the instrumental music. I have also shown that the segmental and morphemic properties of the linguistic interpretation are a by-product of assigning meaning to music through contextual clues and game events.

5. Phonetic aspects of mapping meaning to music

5.1. Methodology

I detail my acoustic investigation on the conventionalised interpretation in this section. For this study, ten native speakers of Yorùbá were recruited in Vancouver, British Columbia. The participants are 4 females and 7 males who immigrated from Nigeria and have spent at most 7 years in Canada. They were all between the ages of 26 and 45. The stimuli in this study are the relevant Super Mario motifs and their conventionalised linguistic interpretations in Yorùbá. Because the sounds in Super Mario Bros often overlap, the relevant Super Mario musical motifs were extracted from YouTube videos of the instrumental music (sources Youcanplayit, 2012a,b,c; Luuul's, 2016). The motifs were performed on a piano by the YouTube sources, unlike the original game music which were performed on a synthesizer.

The original Super Mario music and their rendition in the YouTube video do not include the scores of the music. So, the music scores, even though authentic, are probably approximations of the original videogame music. Any variation between the original Super Mario music, their YouTube rendition and the music score in §3 is probably due to this. For confirmation and replication of the present study, I have included the supplementary material in this paper.

The linguistic interpretations with and without the corresponding instrumental music of Super Mario were presented to each participant. When the participants were presented the linguistic interpretations without the videogame music, they were instructed to read the linguistic interpretations in speech mode. When the linguistic interpretations were presented with the videogame music from the YouTube source, the participants were instructed to recite the phrases in the melody and rhythm of the videogame music. I refer to this recitation as *song* throughout this work. For each participant, each stimulus was repeated three times in speech mode as well as in song mode. The data were recorded with a SHURE WH30XLR cardioid condenser (a headset microphone) at the sampling rate of 48.1 kHz in. wav format. The microphone was attached to a zoomQ8 camera.

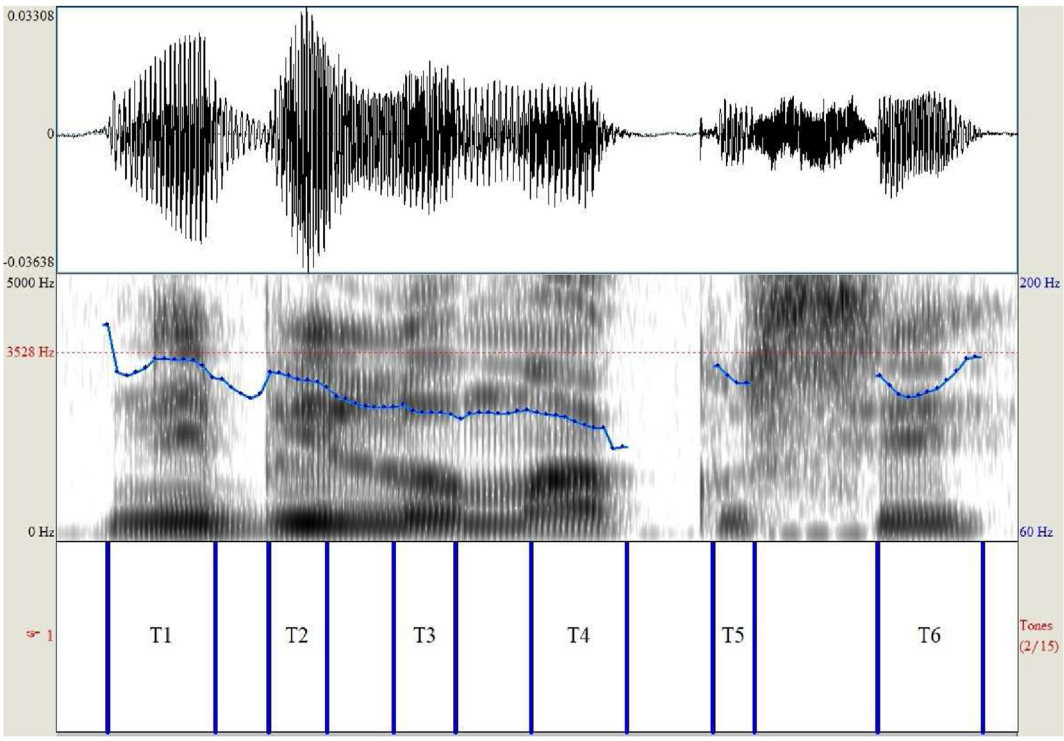

Fig. 6. Waveforms and spectrograms for underworld motif in speech mode. Blue vertical lines show annotation boundaries. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The notes of the relevant Super-Mario music were manually annotated in Praat (Boersma, 2001) and analogously for the corresponding tones of the verbal imitation (i.e., T1, T2, etc.), as shown in Figure 6. To compare the pitch contours of the music motifs and the tones of the corresponding verbal interpretations, F0(Hz) values of the pitch contour were extracted at 25%, 50% and 75% intervals for each music note (and analogously for each tone).

To calculate the correspondence between pitch trajectories of Mario, speech and song sources, I 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 of this statistical measurement). The null hypothesis is that there is no relationship between the musical motifs and the corresponding verbal interpretations in speech and song mode. If the p-value is ≤ 0.05, there is a statistically significant relationship between the pitch contours of the music motifs and those of their corresponding verbal interpretations. Therefore, there is a strong evidence against the null hypothesis. A p-value of > 0.05 indicates weak evidence against null hypothesis. The Pearson's correlation coefficient were calculated using *ggpubr* (Kassambara, 2018) in R (R Core Team, 2022). In the next subsection, I present the results of the production experiment.

5.2. Pitch of Super Mario music and the corresponding Yorùbá interpretations

The F0(Hz) trajectories of the musical motifs and those of their corresponding linguistic interpretations are presented in this section. As mentioned earlier, the relevant Underworld motif has two iterations with different interpretations. For ease of identification, the first and second iterations of the motif are labelled Underworld1 and Underworld2 respectively. The line plots of *Underworld1*, *Underworld2* and the other musical motifs are presented in Figs. 7–11. For each figure, the y-axis contains the F0(Hz) values, and the x-axis contains the normalised time interval for sequences of tone and music-note. The F0(Hz) values are grouped based on the source, namely Mario, speech and song. The results from each participant are grouped by panels (i.e. p1, p2, p3 p#). The results of participant p2 was excluded from the analysis because his Yorùbá accent is heavily influence by Naijá, aka Nigerian Pidgin.

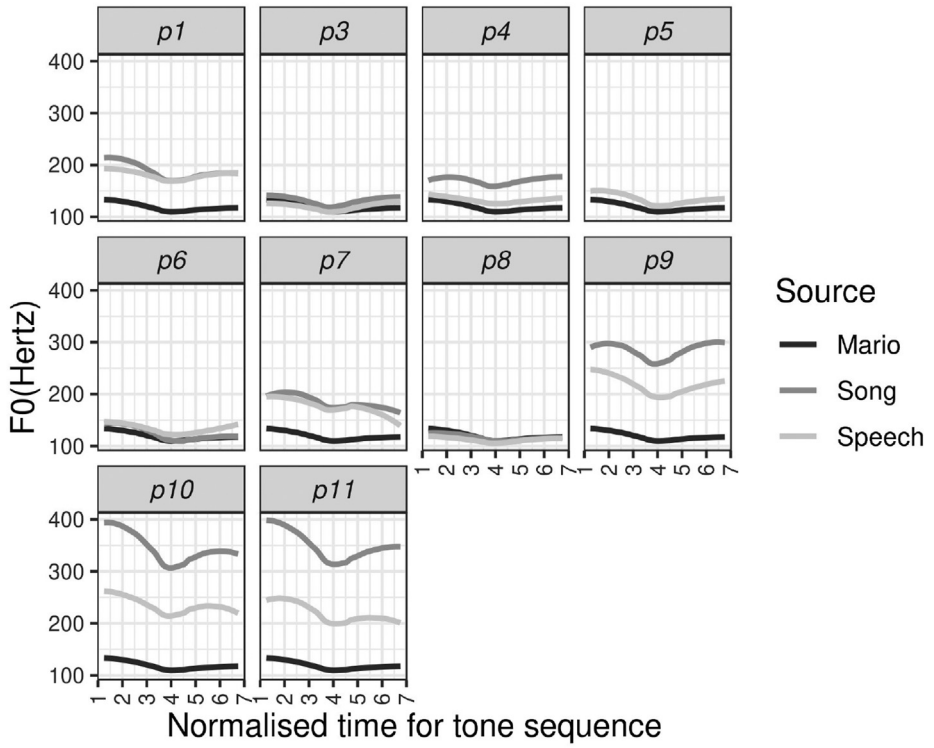

Fig. 7. Pitch contours of Underworld1 in Mario, speech and song.

For Underworld1 in Figure 7, Underworld2 in Figure 8, Overworld in Figure 9 and Death motifs in Figure 10, the pitch contours of the Super Mario motifs and those of their verbal interpretation in song and speech modes are similar at all time intervals. However, in the verbal interpretation of the Overworld motif by participant 11, the pitch contour of the song source does not match that of the corresponding Mario source. The results also show tha the pitch contours of the spoken and sung interpretations are higher than those of the Mario sources for all the participants. When compared to the speech interpretation, the song interpretation has higher pitch. This is expected considering that singing requires more vocal demand than speech.

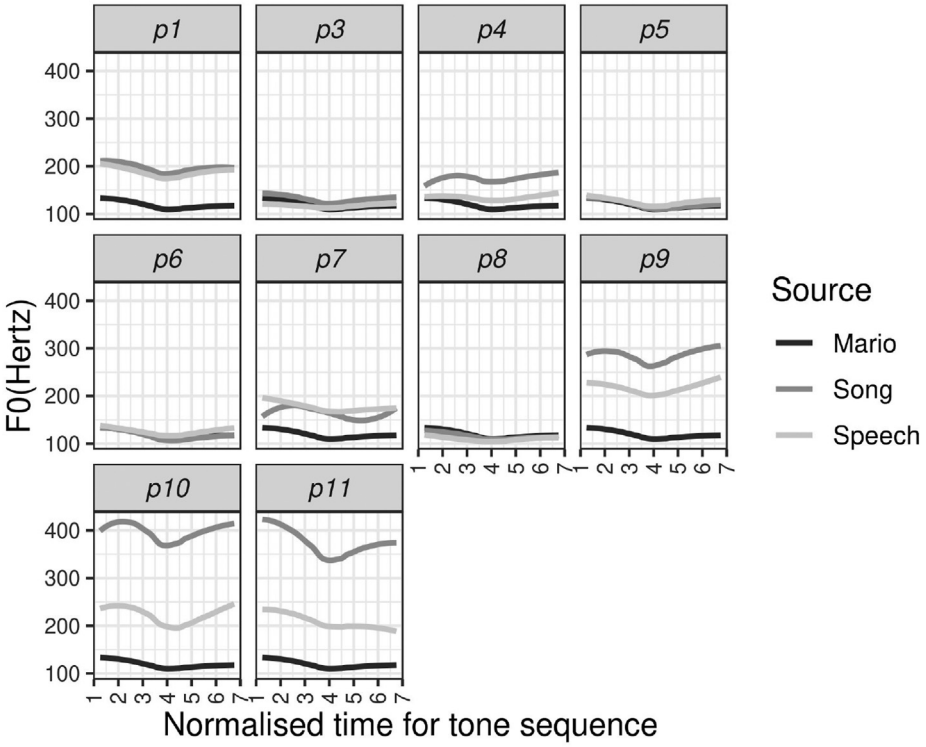

Fig. 8. Pitch contours of Underworld2 in Mario, speech and song.

The pitch contours of the spoken and sung interpretations in Figs. 7 and 8 appear flat when compared to the octave jumps of the Underworld motif score in Figure 4. There are three main factors that might have contributed to the distinction. First, the music scores in §3 are based on the original videogame music, not YouTube source that form the basis of my acoustic investigation. Second, the pitch distinction between the notes of the Underworld motif is smaller when compared to the pitch distinction between the tones of the corresponding spoken and sung interpretations. Third, the target of sound-meaning mapping in speech surrogacy system is pitch contours not pitch height, as shown in Fig. 1. It is worth mentioning that the pitch height of each spoken and sung tones might be conditioned by various linguistic factors. For instance, there is "[an] overall available pitch range for…[every] speaker…and where in the pitch range… [each] tone should be produced" (Yip 2002:11). The pitch range of a tone can also be conditioned by vowel types, the onset of a syllable or the position of the tone-bearing unit (Hombert, 1977; Whalen et al., 1999). For example, the pitch value of a tone is higher when the tone-bearing unit is a high vowel and lower when the tone-bearing unit a low vowel (Whalen and Levitt, 1995).

Unlike the results of the other musical motif, the pitch contour of Flagpole motif and that of the spoken and sung interpretation are only similar at the middle and the end for most participants. For example, at the beginning, the pitch contour of the verbal interpretation rises instead of falling like the Mario source.

With the line plots alone, it is difficult to quantify the strength of the relationship between the pitch contours of the Mario motifs and the corresponding linguistic interpretations. To understand the strength of the relationship between the pitch

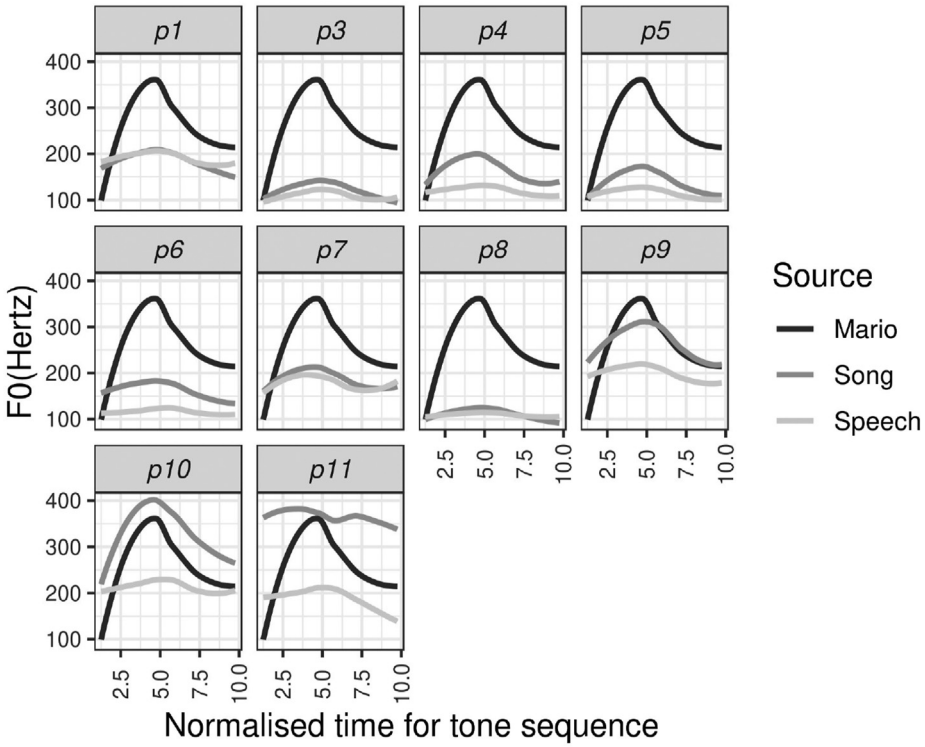

Fig. 9. Pitch contours of Overworld in Mario, speech and song.

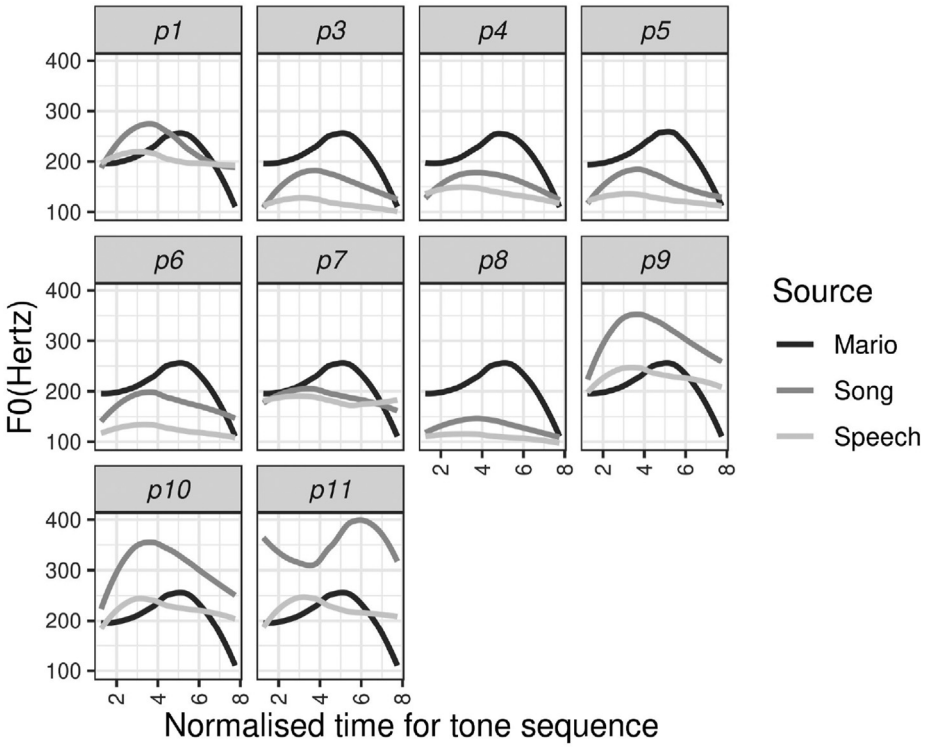

Fig. 10. Pitch contours of Death motif in Mario, speech and song.

contour of the music motifs and the corresponding spoken and sung interpretations, I turn to the results of the Pearson's correlation coefficient. The results are presented in Table 1.

As shown in Table 1, there is a positive linear relationship between the pitch contours of the Mario music motifs and their linguistic interpretations in song mode or speech mode, but the strength of the relationship varies based on music motif and

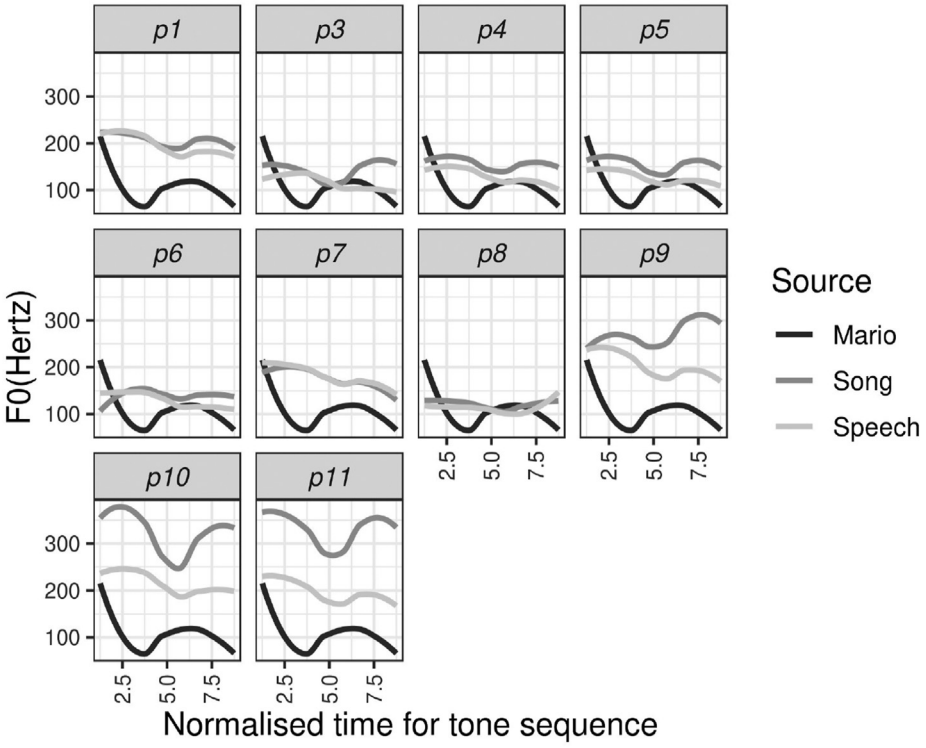

Fig. 11. Pitch contours of Flagpole motif in Mario, speech and song.

mode of verbal interpretation. Regardless of the modes of verbal interpretations, the relationship between the Mario music motifs and their corresponding linguistic interpretations are stronger for Underworld and Overworld motifs ($R \geq 0.58$) than for Death and Flagpole motifs. Relative to the speech renditions, the pitch contours of the song renditions are closely related to those of the Mario instrumentals. With the exception of the Death and Flagpole motifs, the relationships between the pitch contours of all the musical motifs and their spoken and sung interpretations are statistically significant.

The results also show that there is a positive relationship between the pitch contours of the interpretation in speech and song modes, but the degree of similarity varies depending on the music motif which is interpreted. If we take into account that vowel and consonant types affect the pitch value of a tone (Hombert, 1977; Whalen and Levitt, 1995; Whalen

Table 1
Correlation coefficients of Mario motifs and the spoken/sung interpretations.

		SPEECH VS SONG	MARIO VS SPEECH	MARIO VS SONG
Underworld1	*R*	0.98	0.94	0.96
	p	<0.001	<0.001	<0.001
Underworld2	*R*	0.94	0.92	0.93
	p	<0.001	<0.001	<0.001
Overworld	*R*	0.93	0.58	0.75
	p	<0.001	0.001	<0.001
Death	*R*	0.78	0.31	0.7
	p	<0.001	0.172	<0.001
Flagpole	*R*	0.73	0.2	0.13
	p	<0.001	0.34	0.537

et al., 1999), the varying degree of positive correlations between the Mario motifs and their corresponding spoken and sung interpretations can be considered an effect of segmental features on the pitch value of a tone sequences. The effects of segments on the degree of similarity become apparent when we compare the results of Underworld1 and Underworld2. The Underworld1 and Underworld2 having slightly different correlation coefficients is plausibly the effects of the segmental properties of their corresponding sung and spoken interpretations, given that the Underworld1 and Underworld2 have the same sequence of music notes and the same tone sequences but different segments in their sung and spoken interpretations. Put differently, the segmental distinction between Underworld1 and Underworld2 must be the factor that contributes to the varying degree of similarity between the musical motifs and their corresponding spoken and sung interpretations.

The results of the verbal interpretations are similar to the Yorùbá talking drum. As reported in Akinbo (2019, 2021b), Yorùbá drummers communicate by imitating the pitch contours of tones with their native drums and by representing each syllable with a drum strike. The results of the study show that there is a strong positive relation between the pitch contours of Yorùbá tones and those of their corresponding drum renditions. The representation of Yorùbá tones and syllables with a talking drum is consistent sound imitation: "a process by which an individual either vocally or non-vocally generates sounds with qualities that reproduce elements of previously experienced sounds" (Mercado III et al. 2014: 39).

Another interesting aspect of the results involves the correlation between the pitch trajectories of the interpretations in speech and song modes. Although there is a positive correlation between the speech and song of the musical interpretations, the strength of the relationship varies depending on the Mario instrumentals which is musically interpreted. In fact, the correlation between the pitch contour of the speech and song for the Death tune instrumental is not statistically significant. This is in line with the hypothesis that the pitch trajectories of music melody in a tone language is not determined by language, but music can accommodate language when it is musically feasible (Schellenberg, 2012, 2013; McPherson and Ryan, 2018).

The discussion in this section is summarised as follows. Yorùbá gamers linguistically interpreted the Super Mario musical motifs by mapping the pitch contours of the music motifs to Yorùbá phrases with similar pitch contours. Specifically, the gamers achieved this pitch-contour matching by modifying a suitable Yorùbá saying or generating strings of words with similar pitch contours as the music motifs. In this case, the tone sequences of the linguistic interpretations are determined by imitating the pitch contour of the videogame music.

6. Formal analysis and implications of musical meaning in Yorùbá

The goal of this section is to present a formal account of how Yorùbá gamers interpreted the Super Mario musical motifs. For formal accounts of musical meaning, Patel (2008) identifies two approaches, namely semantic and pragmatic. In the semantic approach, the idea is that instrumental music can prime representations of meaningful concepts (Hacohen and Wagner, 1997; Koelsch et al., 2004; Steinbeis and Koelsch, 2011). Under the pragmatic approach, musical meaning is derived through contextual information and multimodal channels (Feld, 1984; Agawu, 2001; Patel, 2008; García et al., 2014; Neumeyer, 2015). To account for the linguistic interpretation of the Super Mario motifs, I adopt the music intepretative moves of Feld (1984), which can be considered a pragmatic approach.

In his work on music communication, Feld (1984) argues that musical structures exist in a social construct, and they have meaning through social interpretations. In this approach, musical meaning is derived from both the internal structure of musical discourse and the situational contexts relating to musical performance or consumption. Feld (1984) proposes the music communication model, which is schematically represented in Fig. 12, as the interpretive move for musical meaning. In the model, music communication involves two dynamically linked components, namely dialectics of sound objects and interpretive moves. The interpretive moves contain five elements: (i) a *locational* element which relates a sound object to an appropriate range within a subjective field of like or unlike items/events; (ii) a *categorical* move which relates sound to things (e.g anthems and patriotic songs); (iii) an *associational* element which relates the sound event to a visual, musical or verbal imagery; (iv) a *reflective element* which relates a sound object to personal and social conditions, and experiences where things like and unlike the object can be heard; (v) an *evaluative* move which involves an affective meaning (Feld, 1984, p. 8). While all or some of the five elements might be utilised in the interpretative moves, the element utilised in an interpretative move must interact with the *identity* of the listener, an *expressive ideology* and *world sense coherence*.

The interpretive moves proposed in Feld (1984) can account for the interpretation of the Super Mario motifs. For this purpose, I refer to the discussion in the previous section. The discussion suggests that the Yorùbá gamers used their cultural experience, vocal imitation and situational contexts of the game in their interpretation of the Super Mario music motifs. Drawing on their native cultural experience in the interpretation of the Super Mario musical motif can be considered a

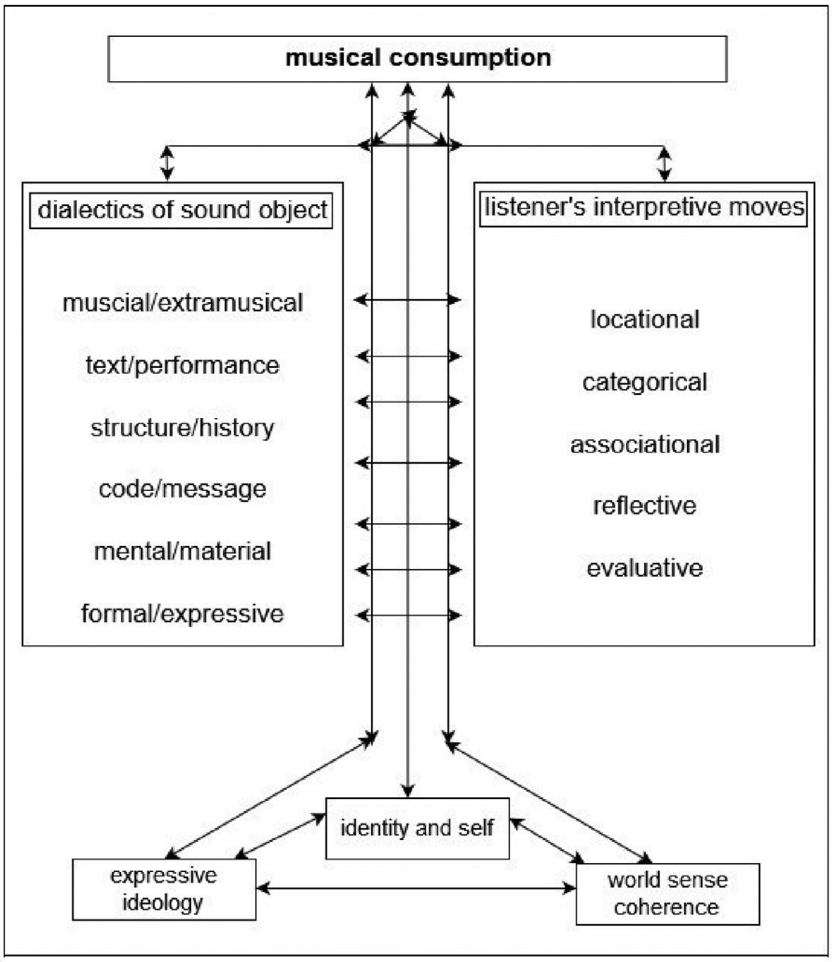

Fig. 12. Interpretive moves (Feld, 1984: 9).

reflective interpretive move. That the gamers imitated the pitch contours of the instrumental music can be likened to a locational interpretive move, which involves relating a sound object to an appropriate range within a subjective field of like or unlike items. In this case, the musical motifs of the game are possibly likened to the talking-drum music in Yorùbá. Given that the associational interpretive move involves relating sound event to a visual, musical or verbal imagery, the use of the visual events in the game for the interpretation of music is consistent with an effect of the associational element.

I have classified the formal account of musical meaning in Feld (1984) as a pragmatic approach, but pragmatic approach to musical meaning, as conceptualised by Patel (2008), solely focuses on the internal structure of musical phrase. Patel (2008) argues for a pragmatic approach as a more efficient way to studying musical meaning. Central to the pragmatic approach in Patel (2008) is the coherent structure of discourse. For instance, the sentences or clauses in (6a) are coherent because they are related in systematic ways and form a unified meaningful whole. In line with the maxim of relevance (also relation) in Grice's Cooperative principle (Grice, 1975; Leech, 1983), the second utterance is relevant because it contributes to the communicative goal.

(7) Linguistic discourse (from Hobbs, 1979)
a. Coherent [John can open Bill's safe]. [He knows the combination].
b. Incoherent [John took a train from Paris to Istanbul].[He likes spinach].

By assuming that coherence relations between clauses in linguistic discourse are analogous to coherence relations between phrases and motifs in musical discourse, Patel (2008) suggests that investigating coherence in musical discourse could offer more valuable insights on musical meaning. The pragmatic approach, as proposed by Patel (2008), predicts that the internal structure of a musical piece, such as the number of musical notes, should have an effect on musical meaning. An argument in support of the pragmatic approach in Patel (2008) is an experiment on the cognition of *raj nplaim* music, which is a musical speech surrogate of Hmong people (Poss, 2012). The results of the experiment shows that the participants were able to match Hmong words to the melodies and rhythm of the raj nplaim instrument, but they found it easier to match phrases to

longer musical melodies. That the participant performed better on longer musical phrases could be an effect of coherence. The prediction of the pragmatic account in Patel (2008) is not completely compatible with the interpretation of Mario music motifs by Yorùbá gamers, given that they also used the context of musical performance in addition to imitating the pitch sequences of the instrumental music.

I now compare the pragmatic approach in this work to the semantic approach which holds that musical features, such as pitch, tempo and rhythm, are capable of priming affective and non-affective meanings (Scherer and Oshinsky, 1977; Kellaris and Kent, 1993; Gagnon and Peretz, 2003; Webster and Weir, 2005; Steinbeis and Koelsch, 2011; Eerola et al., 2013; Dalla Bella et al., 2001; Mote, 2011). While studies adopting semantic approach to musical meaning ignores the role of situational context in musical meaning, the evidence for musical meaning from such studies come from cognitive experiments that involve pairing music with external stimuli, such as pictures and words (e.g., Koelsch et al., 2004: 302).

There is evidence to suggest that music can prime semantic concepts, but as Patel (2008: 334) notes, it does not mean "that music has a semantic system on par with language". For instance, unlike human languages where vocal or visual signals without any external stimuli can have a semantic meaning, an instrumental music without a matching word or any other stimuli has not been shown to elicit a semantic processing (see Hacohen and Wagner, 1997; Steinbeis and Koelsch, 2011; Koelsch, 2005). Of course, the prediction of the semantic approach might not hold in a music tradition with musical speech surrogate like Yorùbá. For instance, a specific pitch value or tone might be associated with words that refer to opposing concepts such as happiness and sadness. If music alone really elicits a semantic processing, the music notes in the Underworld motif of Mario and its repetition would have been assigned the same meaning. Considering that Yorùbá gamers utilised their cultural background and visual information of the game in their interpretation of the background music, it is highly possible that the meaning of the instrumental music in music-word pairing experiments (e.g., Koelsch et al., 2004; Steinbeis and Koelsch, 2011) is derived from coupling music with the visual signals, not music alone. Following from this, the semantic account of musical meaning is not compatible with the interpretation of Mario instrumentals by Yorùbá gamers.

The discussion in this section has shown that the mapping of linguistic meaning to musical sound is better captured by a model that considers internal structure of music phrase, contexts of music performance and cultural background of music consumers.

7. Discussion, summary and conclusion

This study has investigated the linguistic interpretation of Super Mario musical motifs by Yorùbá gamers. The results of the study suggest that the music motifs were interpreted through multimodal channels or moves which might be unordered. The first step involves one-to-one mapping between syllables and music notes of the motifs. To determine the tone of the syllables, the second move involves mapping the pitch contours of the musical motifs to Yorùbá phrases with similar pitch contours. However, the degree of similarity between the pitch contours of the musical motifs and their linguistic interpretations varies depending on the segments of the sung and spoken interpretations. In the third move, the choice of segments or morphemes which is mapped to the music is determined by the situational contexts of the musical performance. In this case, situational contexts include the visual events surrounding the instrumental music in the game, non-virtual events comparable to the visual events in the game and the cultural background of the listeners.

As reported in studies on Yorùbá talking drums, the communicative capability of the speech surrogates is based on mapping the pitch contours of speech forms to the pitch sequences of musical form (Villepastour, 2010, 2014; Akinbo, 2019, 2021b; González and Oludare, 2022). That the gamers vocally imitated the pitch contours of music notes suggests a similarity between musical speech surrogates and linguistic interpretation of music. The strategies used in talking-drum communication and the interpretation of Super Mario music strongly suggest that the mapping of music to linguistic meaning and vice versa involve sound imitation.

The results of this study point to some parallels between sound-meaning mapping in music and language. First, sound-meaning mapping in music and language involves acoustic imitation. For instance, it is well established that the meaning of onomatopoeic words is based on perceptual resemblance between referent and linguistic form (Assaneo et al., 2011; Tsur, 2001; Bezat et al., 2014). While imitation has been recognised as one of the parallels between language and music (Patel, 2008; Jackendoff, 2009), the role of imitation in musical meaning has mostly gone unnoticed. This is probably because studies on musical meaning mostly focus on data from music traditions without speech surrogates and from experimental conditions, not natural contexts of music consumption. Even when research focuses on speech surrogates, the role of imitation in decoding the messages of speech surrogates is only alluded to (Armstrong, 1954; Villepastour, 2010, 2014; Agawu, 2016). This work presents natural evidence for imitation as an strategy for musical meaning in a culture with speech surrogate.

Another parallel between sound-meaning mapping in language and music involves the role of contextual clues. Studies indicate that contextual clues do not only support language acquisition but enhances the retrieval of previous knowledge about sound objects in sound-meaning mapping (Fraser, 1999; Çetinavcı, 2014; Nation, 2015; van den Broek et al., 2018). Similar to sound-meaning association in spoken language, the results of the present study indicate that the interpretation of the musical motifs is context-dependent. That said, there is reason to believe that musical meaning in traditions with speech surrogates may also be context free if we consider that communicative capability of speech surrogates lies in their ability to imitate linguistic features (Akinbo, 2019, 2021b; McPherson, 2019).

The verbal imitation of instrumental music is comparable to warblish, a "non-onomatopoeic verbal mimicry of avian vocalisation" (Sarvasy, 2016, p. 766). Just as the verbal imitation of Super Mario music by Yorùbá gamers, context and imitation play a role in English warblish, but the verbal imitation in English are mostly incoherent. Future research should investigate whether the incoherence of English warblish stems from the fact that, unlike Yorùbá music tradition, speech surrogacy is not a prominent feature of Western music traditions.

That the gamers only interpreted a portion of the game music might be the effect of drum performance and auditory perception on sound-meaning association in music. For example, the drum performance in traditional African music involves switching between purely musical rhythm and speech-like rhythm used in the context of surrogacy (Euba, 1990; Villepastour, 2010; Agawu, 2016). To a large extent, a Yorùbá speaker is aware when a drummer switches from purely musical rhythm to speech-like rhythm and vice versa. This is established in the perceptual experiment of Durojaye et al. (2021a). The assumption therefore is that the uninterpreted portion of the music themes are perceived as music mode of speech surrogacy. The logical move for future research is to compare the rhythm of Yorùbá speech and song to the rhythm of the (un)interpreted portions of the videogame music (as in Patel and Daniele, 2003).

The overlap between sound-meaning mapping in language and music is also crucial to the proposal in Emergent Grammar that language involves general cognitive properties such as memory of phonetic details, paying attention to details and defining similarities across categories (e.g., Hopper, 1987; Bybee, 1998; Archangeli and Pulleyblank, 2022). The fact that language and music have cognitive properties in common is well established in the fields of cognitive science and neuroscience (Steinbeis and Koelsch, 2008; Moreno, 2009; Rebuschat et al., 2011; Patel, 2012; Jungers et al., 2016). For example, in studies by Koelsch and his collaborators (Koelsch et al., 2004; Steinbeis and Koelsch, 2008), they find that "the processes and brain structures involved in the perception of syntax and semantics in music have considerable overlap with those involved in language perception". While most if not all the brain-imagery studies are based on Western music, the findings can be extended to Yorùbá and other cultures with talking drums and other kinds of speech surrogates. That vocal imitation is an aspect of sound-meaning mapping in music and language is not surprising, considering that imitation is a core aspect of general human cognition (Mercado III et al., 2014).

Given that the linguistic interpretations of the music motifs reported here developed naturally and conventionalised with many iterations of generational overturns, there are certain limitations of the present study. For instance, we are uncertain whether, in addition to the factors mentioned in this work, other factors play a role in the linguistic interpretation of the Super Mario music. It would be interesting to investigate whether there is a meaningful distinction between interpreting music on the fly and a conventionalised pattern as is the case here. In a controlled environment, an important issue to address in future studies is how the context of music performance contributes to the linguistic interpretation of instrumental music. For such studies, embedding the musical stimuli as the background music of a specially designed videogame could be promising. Such studies will not only contribute to the fields of linguistics, psychology and music but the emerging field of videogame music, where there is a consensus that background music of a videogame increases the level of "cognitive immersion" or "being in the game" (e.g. Zehnder and Lipscomb, 2006; Grimshaw, 2008; Nacke et al., 2010; Sanders and Cairns, 2010; Zhang and Fu, 2015). The use of cultural experience and situational contexts in the interpretive moves by the Yorùbá gamers is possibly because music in African tradition is strongly functional, linked to dance and externally motivated by social and musical contexts (Agawu, 2001, 2006). Future research should investigate whether functional and non-functional music are interpreted differently.

It is unknown how these interpretations developed. Apart from the fact that the interpretations developed in South-West Nigeria, no one knows the exact time and place where it started. The significant of these interpretations for the gamers has also not being studied. While all these are relevant issues, we might only be able to address a few of these issues in future research.

To conclude, Yorùbá gamers linguistically interpreted instrumental music by vocally imitating the acoustic features of the musical motifs. By considering the immediate contexts of musical performance and the social expectation about music in the relevant contexts, the music consumers assign lexical meaning to the musical motifs. This study suggests some parallels between sound-meaning mapping in language and music. Given the immersive power of videogame, I strongly recommend videogame as an experimental tool for investigating the role of context in music and language studies.

Data availability

Data will be made available on request.

References

Adedeji, Joel A., 1972. The Origin and form of the Yoruba masque theatre. Cah. d'Études Afr. 12 (46), 254–276.
Agawu, V Kofi, 2001. African music as text. Res. Afr. Lit. 32 (2), 8–16.
Agawu, V Kofi, 2006. Structural analysis or cultural analysis? Competing perspectives on the "standard pattern" of West African rhythm. J. Am. Musicol. Soc. 59 (1), 1–46.
Agawu, V Kofi, 2016. The African Imagination in Music. Oxford University Press, Oxford.
Akinbo, Samuel, 2019. Representation of Yorùbá tones by a talking drum: an acoustic analysis. LLA, 11–23.
Akinbo, Samuel, 2021a. Featural affixation and sound symbolism in Fungwa. Phonology 38 (4), 537–569.

Akinbo, Samuel Kayode, 2021b. The Language of Gángan, a Yorùbá Talking Drum 6. https://doi.org/10.3389/fcomm.2021.650382. URL. https://www.frontiersin.org/article/10.3389/fcomm.2021.650382.
Akinbo, Samuel Kayode, 2021c. Vowel Harmony and Some Related Processes in Fungwa. University of British Columbia. Dissertation.
Akintola, Mubarak, 2011. In: Yes, we are the generation that 1st played nintendo, mario (nibi loma kusi - aw we love dat soundtrack) rt. URL. https://twitter.com/AkinMubarak/status/136700690423746561.
Apter, Andrew, 1998. Discourse and its disclosures: Yoruba women and the sanctity of abuse. Africa 68 (1), 68–97.
Archangeli, Diana, Pulleyblank, Douglas, 2022. Emergent Phonology. Language Science Press, Berlin.
Armstrong, Robert G., 1954. Talking drums in the benue-cross river region of Nigeria. Phylon (1940-1956) 15 (4), 355–363.
Assaneo, María Florencia, Nichols, Juan Ignacio, Alberto Trevisan, Marcos, 2011. The anatomy of onomatopoeia. PLoS One 6 (12), e28317.
Auslander, Philip, 2006. Musical personae. TDR/The Drama Review 50 (1), 100–119.
Awóbùlúyì, Oládélé, 1978. Essentials of Yorùbá Grammar. Oxford University Press Nigeria, Ibadan.
Ayoola, Simbiat, 2019. Retro: 7 things Nigerian kids did for fun in the old days. Legit. URL. https://www.legit.ng/1245301-retro-7-nigerian-kids-fun-days.html.
Babel, Molly, Russell, Jamie, 2015. Expectations and speech intelligibility. J. Acoust. Soc. Am. 137 (5), 2823–2833.
Bamgbọṣe, Ayo, 1967. A Short Yorùbá Grammar. Heinemann, Ibadan.
Bezat, Marie-Céline, Kronland-Martinet, Richard, Vincent, Roussarie, Ystad, Sølvi, 2014. From acoustic descriptors to evoked quality of car door sounds. J. Acoust. Soc. Am. 136 (1), 226–241.
Blench, Roger, 2019. An Atlas of Nigerian Languages. Kay Williamson Educational Foundation, Cambridge.
Boersma, Paul, 2001. Praat, a system for doing phonetics by computer. Glot Int. 5 (9–10), 341–345.
Bradley, David, 1979. Speech through music: the Sino-Tibetan gourd reed-organ. Bull. Sch. Orient Afr. Stud., 535–540. University of London.
van den Broek, Gesa SE., Takashima, Atsuko, Segers, Eliane, Verhoeven, Ludo, 2018. Contextual richness and word learning: context enhances comprehension but retrieval enhances retention. Lang. Learn. 68 (2), 546–585.
Bybee, Joan, 1998. The emergent lexicon. In: Catherine Gruber, M., Higgins, Derrick, Olson, Kenneth S., Wysocki, Tamra (Eds.), Proceedings of the Chicago Linguistic Society CLS 34, vol. 34. Chicago Linguistic Society, Chicago, Illinois, pp. 421–435.
Campbell, Bolaji, 2015. Eegun ogun: war masquerades in Ibadan in the era of modernization. Afr. Arts 48 (1), 42–53.
Carrington, John F., 1971. The talking drums of Africa. Sci. Am. 225 (6), 90–95.
Çetinavcı, Berrin Manga, 2014. Contextual factors in guessing word meaning from context in a foreign language. Procedia-Social and Behavioral Sciences 116, 2670–2674.
Cochrane, Tom, 2010. Using the persona to express complex emotions in music. Music Anal. 29 (1–3), 264–275.
Cohen, Annabel J., 2011. Music as a source of emotion in film. In: Juslin, Patrik N., Sloboda, John (Eds.), Handbook of Music and Emotion: Theory, Research, Applications. Oxford University Press, Oxford, pp. 879–908.
Collins, Karen, 2009. An Introduction to procedural music in video games. Contemp. Music Rev. 28 (1), 5–15.
Dalla Bella, Simone, Peretz, Isabelle, Rousseau, Luc, Gosselin, Nathalie, 2001. A developmental study of the affective value of tempo and mode in music. Cognition 80 (3), B1–B10.
Devos, Thierry, Banaji, Mahzarin R., 2005. American= white? J. Personality Soc. Psychol. 88 (3), 447.
Dingemanse, Mark, Blasi, Damián E., Gary, Lupyan, Christiansen, Morten H., Monaghan, Padraic, 2015. Arbitrariness, iconicity, and systematicity in language. Trends Cognit. Sci. 19 (10), 603–615.
Durojaye, Cecilia, Fink, Lauren, Roeske, Tina, Wald-Fuhrmann, Melanie, Larrouy-Maestri, Pauline, 2021a. Perception of Nigerian dùndún talking drum performances as speech-like vs. music-like: the role of familiarity and acoustic cues. Front. Psychol. 12, 1760.
Durojaye, Cecilia, Knowles, Kristina L., Jakob Patten, K., Garcia, Mordecai J., McBeath, Michael K., 2021b. When music speaks: an acoustic study of the speech surrogacy of the Nigerian dùndùn talking drum. Frontiers in Communication 132.
Eberhard, Kathleen M., Spivey-Knowlton, Michael J., Sedivy, Julie C., Tanenhaus, Michael K., 1995. Eye movements as a window into real-time spoken language comprehension in natural contexts. J. Psycholinguist. Res. 24 (6), 409–436.
Eerola, Tuomas, Friberg, Anders, Bresin, Roberto, 2013. Emotional expression in music: contribution, linearity, and additivity of primary musical cues. Front. Psychol. 4, 487.
Ekwueme, Lazarus N., 1974. Linguistic determinants of some Igbo musical properties. J. Afr. Stud. 1 (3), 335–353.
Euba, Akin, 1990. Yorùbá drumming: the Dùndún tradition. Bayreuth African Studies, Bayreuth.
Fairchild, Charles, Marshall, P David, 2019. Music and Persona: An Introduction.
Fámúle, Oláwolé, 2018. In: èdè àyàn: the language of Àyàn in Yorùbá art and ritual of Egúngún. vol. 2. Yoruba Studies Review 2. URL. https://news.clas.ufl.edu/ede-yan-the-language-of-yan-in-yoruba-art-and-ritual-of-egungun/.
Feld, Steven, 1984. Communication, music, and speech about music. Yearbk. Tradit. Music 16, 1–18.
Fraser, Carol A., 1999. Lexical processing strategy use and vocabulary learning through reading. Stud. Sec. Lang. Acquis. 21 (2), 225–241.
Zhang, Jiulin, Fu, Xiaoqing, 2015. The influence of background music of video games on immersion. J. Psychol. Psychother. 5 (4).
Gagnon, Lise, Peretz, Isabelle, 2003. Mode and tempo relative contributions to “happy-sad” judgements in equitone melodies. Cognit. Emot. 17 (1), 25–40.
García, Ricardo R., Zamorano, Francisco, Aboitiz, Francisco, 2014. From imitation to meaning: circuit plasticity and the acquisition of a conventionalized semantics. Front. Hum. Neurosci. 8, 605.
González, Mariano, Oludare, Olupemi Ezekiel, 2022. The speech surrogacy systems of the Yoruba dùndún and bàtá drums. on the interface between organology and phonology. Frontiers in Communication 245.
Green, Thomas, 2005. Dambe: traditional Nigerian boxing. Journal of Combative Sport.
Gregory, Andrew H., Varney, Nicholas, 1996. Cross-cultural comparisons in the affective response to music. Psychol. Music 24 (1), 47–52.
Grice, Herbert P., 1975. Logic and conversation 3: speech acts. In: Cole, Peter, Morgan, Jerry L. (Eds.), Speech Acts. Academic Press, New York, pp. 41–58.
Grimshaw, Mark, 2008. Sound and immersion in the first-person shooter. Int. J. Interdiscipl. Global Stud. 5 (1), 119–124.
Hacohen, Ruth, Wagner, Naphtali, 1997. The Communicative force of Wagner’s leitmotifs: complementary relationships between their connotations and denotations. Music Perception 14 (4), 445–475.
Hasegawa, Koichi, Shinohara, Chika, Broadbent, Jeffrey P., 2007. The effects of ‘social expectation’on the development of civil society in Japan. J. Civ. Soc. 3 (2), 179–203.
Hobbs, Jerry R., 1979. Coherence and coreference. Cognit. Sci. 3 (1), 67–90.
Hockett, Charles F., 1960. The origin of speech. Sci. Am. 203 (3), 88–97.
Hombert, Jean-Marie, 1977. Consonant types, vowel height and tone in Yoruba. Stud. Afr. Ling. 33, 173–190.
Hopper, Paul, 1987. Emergent grammar. In: Aske, Jon, Beery, Natasha, Michaelis, Laura, Hana, Filip (Eds.), Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society, vol. 13. Berkley Linguistics Societry Inc, Berkley, California, pp. 139–157.
Hughes, David W., 2000. No nonsense: the logic and power of acoustic-iconic mnemonic systems. Br. J. Ethnomusicol. 9 (2), 93–120.
Hyman, Larry M., 2018. What tone teaches us about language. Language 94 (3), 698–709.
Isola, Akinwumi, 1982. Ena: code-talking in Yoruba. J. W. Afr. Lang. 12 (1), 43–51.
Jackendoff, Ray, 2009. Parallels and nonparallels between language and music. Music perception 26 (3), 195–204.
James, Lucas, 2021. Systems of communication: aspects of culture and structure in speech surrogates. Frontiers in Communication 6. https://doi.org/10.3389/fcomm.2021.653268. URL. https://www.frontiersin.org/article/10.3389/fcomm.2021.653268.
Jungers, Melissa K., Hupp, Julie M., Dickerson, Sara D., 2016. Language priming by music and speech: evidence of a shared processing mechanism. Music Perception 34 (1), 33–39.

Kang, Okim, Rubin, Donald L., 2009. Reverse linguistic stereotyping: measuring the effect of listener expectations on speech evaluation. J. Lang. Soc. Psychol. 28 (4), 441–456.
Kassambara, Alboukadel, 2018. ggpubr:"ggplot2" based publication ready plots. R package version 0.1.7. URL. https://CRAN.R-project.org/package=ggpubr.
Keil, Charles, Feld, Steven, 1994. Music Grooves: Essays and Dialogues. University of Chicago Press, Chicago.
Kellaris, James J., Kent, Robert J., 1993. An exploratory investigation of responses elicited by music varying in tempo, tonality, and texture. J. Consum. Psychol. 2 (4), 381–401.
Koelsch, Stefan, 2005. Neural substrates of processing syntax and semantics in music. Curr. Opin. Neurobiol. 15, 207–212.
Koelsch, Stefan, Kasper, Elisabeth, Sammler, Daniela, Schulze, Katrin, Gunter, Thomas, Friederici, Angela D., 2004. Music, language and meaning: brain signatures of semantic processing. Nat. Neurosci. 7 (3), 302–307.
Laniran, Yetunde O., Clements, George N., 2003. Downstep and high raising: interacting factors in Yoruba tone production. J. Phonetics 31 (2), 203–250.
Laroche, Guillaume, 2012. Analyzing Musical Mario-Media: Variations in the Music of Super Mario Video Games. MA thesis.
Lee, Martin J., 1985. From rivalry to hostility among sports fans. Quest 37 (1), 38–49.
Leech, Geoffrey N., 1983. Principles of Pragmatics. Longman, London.
Lerner, Neil, 2014. Mario's dynamic leaps: musical innovations (and the specter of early cinema) in Donkey Kong and Super Mario Bros. In: Donnelly, Kevin J. , Gibbons, William, Lerner, Neil (Eds.), Music in Video Games. Routledge, New York, pp. 15–43.
Lo-Bamijoko, Joy Nwosu, 1987. Classification of Igbo musical instruments, Nigeria. Journal of the International Library of African Music 6 (4), 19–41.
Loguidice, Bill, Barton, Matt, 2012. Vintage Games: An Insider Look at the History of Grand Theft Auto, Super Mario, and the Most Influential Games of All Time. Elsevier Focal Press, Oxford.
Luuul's, 2016. In: Super mario Bros – underground theme piano tutorial. URL. https://www.youtube.com/watch?v=0iAlTae3poc.
McGowan, Kevin B., 2015. Social expectation improves speech perception in noise. Lang. Speech 58 (4), 502–521.
McLeod, Ken, 2006. "We are the champions": masculinities, sports and popular music. Popular Music Soc. 29 (5), 531–547.
McPherson, Laura, 2018. The Talking balafon of the Sambla: grammatical principles and documentary implications. Anthropol. Ling. 60 (3), 255–294.
McPherson, Laura, 2019. Musical adaptation as phonological evidence: case studies from textsetting, rhyme, and musical surrogates. Language and Linguistics Compass 13 (12), 1–15.
McPherson, Laura, Ryan, Kevin M., 2018. Tone-tune association in tommo so (dogon) folk songs. Language 94 (1), 119–156.
Mercado III, Eduardo, Mantell, James T., Pfordresher, Peter Q., 2014. Imitating sounds: a cognitive approach to understanding vocal imitation. Comparative Cognition & Behavior Reviews 9.
Moreno, Sylvain, 2009. Can music influence language and cognition? Contemp. Music Rev. 28 (3), 329–345.
Mote, Jasmine, 2011. The effects of tempo and familiarity on children's affective interpretation of music. Emotion 11 (3), 618.
Mullins, Catherine E., 2014. Blah, blah, blah: making sense of nonsense in Irish vocal music. Musical Offerings 5 (2), 87–117.
Munday, Rod, 2007. Music in video games. In: Sexton, Jamie (Ed.), Music, Sound and Multimedia: From the Live to the Virtual. Edinburgh University Press, pp. 51–67.
Nacke, Lennart E., Grimshaw, Mark N., Lindley, Craig A., 2010. More than a feeling: measurement of sonic user experience and psychophysiology in a first-person shooter game. Interact. Comput. 22 (5), 336–343.
Nation, Paul, 2015. Principles guiding vocabulary learning through extensive reading. Read. Foreign Lang. 27, 136—–145.
Neumeyer, David P., 2015. Meaning and Interpretation of Music in Cinema. Indiana University Press, Bloomington & Indianapolis.
Nintendo, Entertainment System, 1985. Super Mario Bros.: Instructional Booklet. Nintendo Co., Ltd. Video Game.
Nketia, J H Kwabena, 1963. Drumming in Akan Communities of Ghana. Thomas Nelson and sons, Edinburgh.
Odesanya, _OPj4, 2013. Who knows that sound in super mario..we yoruba boys translated it to 'ibi lo ma ku si, boya lo ma dele:'. Twitter. URL. https://twitter.com/_OPj/status/306884368075468802.
Omojola, Bode, 2011. Òṣogbo: power, song and performance in a Yoruba festival. Ethnomusicol. Forum 20 (1), 79–106.
Parrott, W Gerrod, Hertel, Paula, 1999. Research methods in cognition and emotion. In: Dalgleish, Tim, Power, Mick (Eds.), Handbook of Cognition and emotion. John Wiley & Sons, West Sussex, pp. 61–81.
Patel, Aniruddh D., 2008. Music, Language, and the Brain. Oxford University Press, New York.
Patel, Aniruddh D., 2012. Language, music, and the brain: a resource-sharing framework. In: Rebuschat, Patrick, Rohrmeier, Martin, Hawkins, John A., Cross, Ian (Eds.), Language and Music as Cognitive Systems. Oxford University Press, Oxford, pp. 204–223.
Patel, Aniruddh D., Daniele, Joseph R., 2003. An empirical comparison of rhythm in language and music. Cognition 87 (1), B35–B45.
Patel, Aniruddh D., Iversen, John R., 2003. Acoustic and perceptual comparison of speech and drum sounds in the North Indian Tabla tradition: an empirical study of sound symbolism. In: Solé, Maria-Josep, Recasens, Daniel, Romero, Joaquin (Eds.), Proceedings of the 15th International Congress of Phonetic Sciences (ICPHS). Causal Productions Pty Ltd, Adelaide, pp. 925–928.
Poss, Nicholas Frederick, 2005. The Communication of verbal content on the Hmong Raj: an ethnographic analysis of performance practice. The Ohio State University. MA Thesis.
Poss, Nicholas Frederick, 2012. Hmong Music and Language Cognition: An Interdisciplinary Investigation. The Ohio State University. Phd Thesis.
Pulleyblank, Douglas., 2004. A note on tonal markedness in Yoruba. Phonology 21 (3), 409–425. https://doi.org/10.1017/S0952675704000326.
Ramachandran, Vilayanur S., Hubbard, Edward M., 2001. Synaesthesia–a window into perception, thought and language. J. Conscious. Stud. 8 (12), 3–34.
Rebuschat, Patrick, Rohrmeier, Martin, Hawkins, John A., Cross, Ian, 2011. Language and Music as Cognitive Systems. Oxford University Press, Oxford.
Rubin, Donald L., 1992. Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. Res. High. Educ. 33 (4), 511–531.
Rumsey, Deborah J., 2009. Statistics essentials for dummies. John Wiley & Sons, New Jersey.
Sanders, Timothy, Cairns, Paul, 2010. Time perception, immersion and music in videogames. Proceedings of HCI 2010 24, 160–167.
Sarvasy, Hannah, 2016. Warblish: verbal mimicry of birdsong. J. Ethnobiol. 36 (4), 765–782.
de Saussure, Ferdinand, 1974. Course in General Linguistics. Fontana, London.
Schartmann, Andrew, 2015. Koji Kondo's Super Mario Bros. Soundtrack. Bloomsbury Publishing, New York.
Schellenberg, Murray, 2012. Does language determine music in tone languages? Ethnomusicology 56 (2), 266–278.
Schellenberg, Murray Henry, 2013. The realization of tone in singing in Cantonese and Mandarin. University of British Columbia. PhD Thesis.
Scherer, Klaus R., Oshinsky, James S., 1977. Cue utilization in emotion attribution from auditory stimuli. Motiv. Emot. 1 (4), 331–346.
Seifart, Frank, Meyer, Julien, Grawunder, Sven, Dentel, Laure, 2018. Reducing language to rhythm: amazonian Bora drummed language exploits speech rhythm for long-distance communication. R. Soc. Open Sci. 5 (4), 170354.
Sotunsa, Mobolanle Ebunoluwa, 2021. Drum language and literature. In: Akinyemi, Akintunde, Falola, Toyin (Eds.), The Palgrave Handbook of African Oral Traditions and folklore. Springer Nature, Cham, pp. 281–295.
Steinbeis, Nikolaus, Koelsch, Stefan, 2008. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. Cerebr. Cortex 18 (5), 1169–1178.
Steinbeis, Nikolaus, Koelsch, Stefan, 2011. Affective priming effects of musical sounds on the processing of word meaning. J. Cognit. Neurosci. 23 (3), 604–621.
Stern, Theodore, 1957. Drum and whistle 'languages': an analysis of speech surrogates. Am. Anthropol. 59, 487–506.
Stuart, Keith, 2010. In: Super mario Bros: 25 mario facts for the 25th anniversary. URL. https://amp.theguardian.com/technology/gamesblog/2010/sep/13/games-gameculture.
Tanenhaus, Michael K., Trueswell, John C., 1995. Sentence comprehension. In: Miller, Joanne L., Eimas, Peter D. (Eds.), Speech, Language and Communication. Academic Press, San Diego, CA, pp. 217–262.

Team, R Core., 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Thackway, Melissa, 2002. Integrating orality into francophone West African film. In: Tobias, Döring (Ed.), African Cultures, Visual Arts, and the Museum: Sights/sites of Creativity and Conflict. Rodopi, Amsterdam/New York, pp. 229–242.
Tsur, Reuven, 2001. Onomatopoeia: cuckoo-language and tick-tocking. Iconicity in language. URL. http://www.trismegistos.com/IconicityInLanguage/Articles/Tsur/default.html.
Vale, Leonor, Fernandes, Teresa, 2018. Social media and sports: driving fan engagement with football clubs on Facebook. J. Strat. Market. 26 (1), 37–55.
Valsiner, Jaan, 1994. What is "natural" about "natural contexts"? Cultural construction of human development (and its study). Infancia Aprendiz. 17 (66), 11–19.
Villepastour, Amanda, 2010. Ancient Text Messages of the Yorùbá Bàtá Drum: Cracking the Code. Ashgate Publishing Limited, Surrey, England.
Villepastour, Amanda, 2014. Talking tones and singing speech among the Yorùbá of Southwest Nigeria. In: Lechleitner, Gerda, Liebl, Christian (Eds.), Jahrbuch des phonogrammarchivs der österreichischen akademie der wissenschaften, vol. 44. Cuvillier, Göttingen, pp. 29–47.
Wagner, Richard, 1964. Wagner on Music and Drama. Da Capo, New York.
Walker, Robert, 1996. Open peer commentary: can we understand the music of another culture? Psychol. Music 24 (2), 103–114.
Waterman, Christopher A., 2000. Yoruba popular music. In: Stone, Ruth M. (Ed.), The Garland Handbook of African Music. Routledge, Oxford, pp. 169–185.
Webster, Gregory D., Weir, Catherine G., 2005. Emotional responses to music: interactive effects of mode, texture, and tempo. Motiv. Emot. 29 (1), 19–39.
Weir, Michele, 2015. The Scat singing dialect: an introduction to vocal improvisation. Choral J. 55 (11), 28.
Whalen, Doug H., Gick, Bryan, Kumada, Masanobu, Honda, Kiyoshi, 1999. Cricothyroid activity in high and low vowels: exploring the automaticity of intrinsic f0. J. Phonetics 27 (2), 125–142.
Whalen, Douglas H., Levitt, Andrea G., 1995. The universality of intrinsic F0 of vowels. J. Phonetics 23 (3), 349–366.
Whalen, Zach, 2004. Play along-an approach to videogame music. Game Stud. 4 (1), 1–28.
Winter, Yoad, 2014. On the grammar of a Senegalese drum language. Language 90 (3), 644–668.
Yi, Han-Gyol, Phelps, Jasmine EB., Smiljanic, Rajka, Chandrasekaran, Bharath, 2013. Reduced efficiency of audiovisual integration for nonnative speech. J. Acoust. Soc. Am. 134 (5), EL387–EL393.
Yip, Moira, 2002. Tone. Cambridge University Press, Cambridge.
Youcanplayit, 2012a. In: Super mario Bros – Flagpole [easy piano tutorial]. URL. https://www.youtube.com/watch?v=ist0vB_idcs.
Youcanplayit, 2012b. In: Super Mario Bros–death–easy piano tutorial. URL. https://www.youtube.com/watch?v=cO8xgXcm_hg.
Youcanplayit, 2012c. In: Super Mario Bros–piano tutorial–easy. URL. https://www.youtube.com/watch?v=eQ_VLqJC8UY.
Zehnder, Sean M., Lipscomb, Scott D., 2006. The Role of music in video games. In: Vorderer, Peter, Bryant, Jennings (Eds.), Playing Video Games – Motives, Responses, and Consequences. Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 241–258.