

# Vocal Imitation and Linguistic Processing of Super Mario Theme Music by Yorùbá Gamers

*Samuel Akinbo*  
*University of British Columbia*

## Abstract

An aspect of gaming culture among Yorùbá millennials is the linguistic interpretations of the background music that accompanies the popular video game called Super Mario. The themes of the linguistic interpretations are comparable to those of music texts at traditional Yorùbá competitions (Apter, 1998; Fámúle, 2018). Drawing on the Yorùbá music tradition, the account in this work is that the gamers assumed that the background music of the video game 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 video game. The results of an acoustic analysis show that the choice of words in the linguistic interpretations are also determined by mapping the pitch trajectories of the music melodies to the tones of the gamers' native language. Notably, the results of this study suggest that the linguistic processing of music may not only involve phonetic iconicity (Hacohen and Wagner, 1997; Koelsch et al., 2004; Steinbeis and Koelsch, 2011) but situational context and social expectation (Feld, 1984; Agawu, 2001; Patel, 2008; García et al., 2014; Neumeier, 2015).

## 1. Introduction

Ancient philosophers and thinkers have contemplated the relations between language and music. For example, the Greek philosopher Plato claimed that the power of music stems from its resemblance to the sounds of speech (Neubauer 1986, cited in Patel 2008). Most strikingly, the music-language connection is embodied in the capability of the talking drums or flutes in many African music traditions. The most popular example of talking drum comes from Yorùbá which is a Benue-Congo language with more than 20 million speakers in South-Western Nigeria (Pulleyblank, 2009). Like most Nigerian languages, Yorùbá is a tone language, which means pitch contrasts bring about lexical or grammatical distinctions in meaning (Yip, 2002). The language has three contrastive tones: H(igh) bá 'meet', L(ow) bà 'land', and M(id) ba 'to plait' (Bamgboṣe, 1967; Awóbùlúyì, 1978; Akinlabi, 1985; Pulleyblank, 1986). Yorùbá musicians communicate through a family of Yorùbá drums called dùndún by varying the compression of the drum's tension cords. The varying compressions of the tension cords encodes the pitch contours of tones in Yorùbá words (Beier, 1954; Euba, 1990; Waterman, 2000; Villepastour, 2010, 2014). Using acoustic analysis, Akinbo (2019) shows that there is a positive correlation between the pitch contours of the speech tones and those of their corresponding musical rendition (Figure 1).

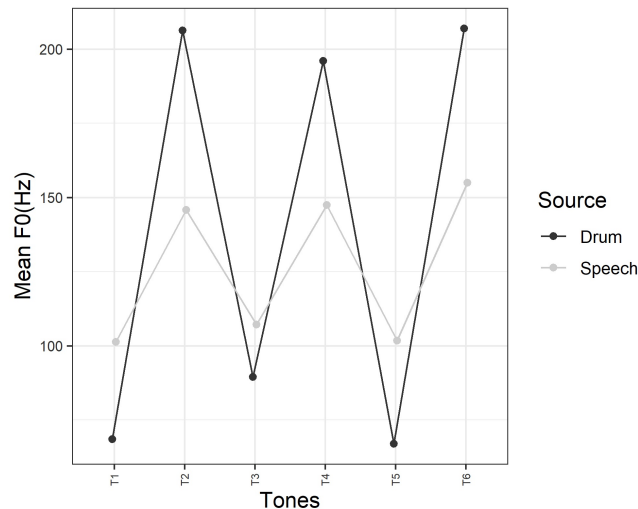


Figure 1: Mean  $F_0$  contours for Yorùbá drum and speech tones

Figure 1 shows the drummed and spoken forms of the Yorùbá phrase [ògúnbòdé gbàgbá] “Ògúnbòdé took garden egg”. In the graph, the y-axis contains the acoustic measurement of pitch contour in  $F_0$ (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 trajectory.

Just as in Yorùbá, the musicians in other cultures around the world are also able to musically communicate by encoding meaning or the phonemic aspects of their languages with music instruments (Stern, 1957; Carrington, 1971; Ekwueme, 1974; Bradley, 1979; Lo-Bamijoko, 1987; Poss, 2005; Winter, 2014; Poss, 2012; Seifart et al., 2018; McPherson, 2018). Studies on 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 language imitation or interpretation of music melodies.

Musicologists and cognitive psychologists have long been investigating how musical melodies are linguistically interpreted (Feld, 1984; Hachohen and Wagner, 1997; Koelsch et al., 2004; Patel, 2008; García et al., 2014; Neumeyer, 2015), but most of these studies focused on music traditions without speech surrogate systems. In a culture with musical speech surrogates, studying musical meaning from a linguistic perspective could offer valuable insights into the role of the acoustic parallel between language and instrumental music in music cognition, and the connection between musical and linguistic communication. Using linguistic instrumentation and methodology, the present study contributes to research on musical meaning. Specifically, this study explores how Yorùbá speakers linguistically interpret unfamiliar instrumental music.

The present work is based on conventionalised interpretations of the background music in the Super Mario game by Yorùbá gamers. The linguistic interpretations are discussed in §2. The discussion in §3 explores why the gamers interpreted the video game music

into Yorùbá. To compare the acoustics of the linguistic interpretation to the acoustics of the corresponding background music, an acoustic investigation was conducted. Details of the acoustic investigation are presented in §4. The results of the acoustic investigation are presented in §5. The role of situational contexts in the linguistic interpretation is explored in §6. In §7, the interpretative moves are analysed within the pragmatic theory of musical meaning (Feld, 1984; Patel, 2008). The summary and the conclusion are presented in §8.

## 2. Interpreting the theme music of Super Mario Bros.

The data source for the present study is the linguistic interpretation of Super Mario music. In this section, I present background information about the game, the theme music of the game and the linguistic interpretation of the theme music by Yorùbá gamers.

Nintendo, which is a multinational Japanese company, released the console video game called Super Mario Bros in 1985. More than 40 million copies of the game were sold worldwide (Stuart, 2010). The protagonist of the 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 theme music of the game. According to Laroche (2012), the composition of the theme music 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. Adaptive audio events 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, 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 theme (see Lerner, 2014; Schartmann, 2015). The musical experience in video games are the major innovations of Super Mario Bros (Collins, 2009; Lerner, 2014).

It is uncertain when the game got to Nigeria, but the game was very popular in Nigeria around 1990s and early 2000s. An aspect of gaming culture among Yorùbá gamers in Nigeria is the linguistic interpretation of Super Mario instrumentals into Yorùbá (Ayoola, 2019). The linguistic 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. The conventionalised interpretation of the Overworld theme is presented in Figure 2.

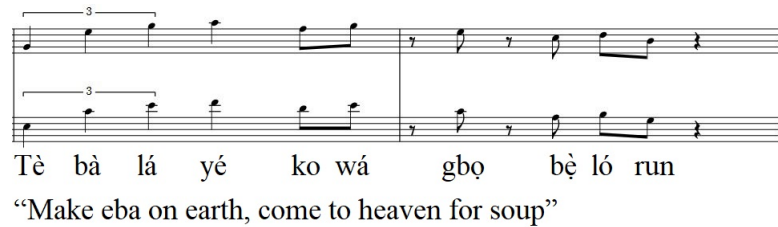


Figure 2: The interpretation of Overworld theme by Yorùbá gamers

The Overworld is played at the first stage of the game. As a result of this, it is also called the main theme. The linguistic interpretation of the Overworld theme is an adaption of a Yorùbá idiom which is used as a death threat. The only difference is that the phrase /gún ’yán/ ‘pound iyán’ in the idiom is replaced with the phrase /tè ’bà/ ‘make èbà (lit: mash èbà)’ in the interpretation of the Overworld theme<sup>1</sup>. With the exception of the Overworld theme, the interpretations of the other theme music were originally created by the gamers.

When Mario is about to enter the castle after winning the first stage of the game, the Flagpole fanfare is played. The linguistic interpretation of the flagpole fanfare by Yorùbá gamers is presented in Figure 3.

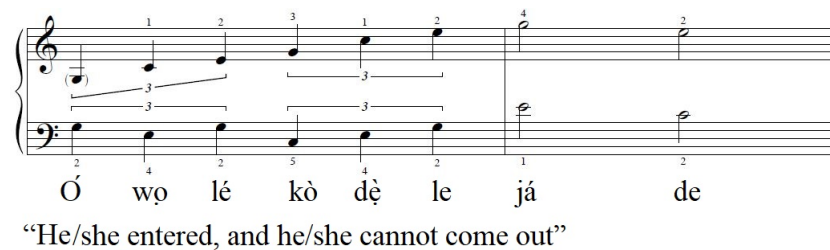


Figure 3: The interpretation of Flagpole fanfare theme by Yorùbá gamers

The second stage of Super Mario Bros is set in the underworld. Immediately Mario gets into the underworld, the Underworld theme is played. The first six notes of the Underworld music theme are repeated once, but the notes and their repetition have different linguistic interpretations. The linguistic interpretations of the Underworld rhythms are presented in Figure 4.

<sup>1</sup>Iyán is the Yorùbá name for pounded yam; èbà is a staple food made from dried grated cassava (manioc).

Ní bí lo ma kú sí, bó yá lo ma dé lé  
 It is here that you are going to die; I don't think you will get home

Figure 4: The interpretation of Underworld theme 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 this theme music and the linguistic interpretation is presented in Figure 5.

E<sup>4</sup> ré<sup>4</sup> ké<sup>4</sup> ré<sup>3</sup> lo<sup>2</sup> wá<sup>1</sup> se<sup>2</sup>  
 "it is a dangerous play that you came to play"

Figure 5: The interpretation of Death theme by Yorùbá gamers

In general, Yorùbá listeners (i.e Super Mario gamers) are able to linguistically translate or interpret certain portions of some theme music in Super Mario Bros. The interpretations are usually performed as a cappella. Every Yorùbá gamer who has played Super Mario knows these conventionalised interpretations, but it is uncertain whether someone who has not heard the theme music before might have the same interpretations.

An interesting feature of the linguistic interpretations is that a lexical syllable is mapped to each music note of the theme music. However, the choice of tone on a syllable is unpredictable. So, it will be interesting to investigate the factors that determine the choice of tone on the syllables. Although studies on 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), it remains to be seen if these features are used in the interpretation of the theme music in Super Mario Bros. These issues are addressed in §4. Before focusing on the tone of the syllables, the next section addresses why the gamers assume that the background music of the game are communicative.

### 3. Why linguistically interpret the theme music?

The discussion so far has shown that the gamers linguistically interpreted the theme music of Super Mario by mapping lexical syllables to the music notes. Instead of mapping lexical syllables to the music melodies of the theme music in Super Mario Bros, the Yorùbá gamers

could have mapped non-lexical vocables or nonsense syllables to the music themes, as is done in other cultures (see Hughes, 2000; Mullins, 2014; Weir, 2015). So, why did the gamers use lexical syllables instead of vocables? In other words, why do the players assume that an unfamiliar music involves linguistic communication? To answer this question, we must understand the role of music in Yorùbá traditional games and events.

In Yorùbá culture, as in other African cultures, music 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 Oròyèyè festival in Ayédé, Èkitì state, Nigeria. The background music is played by the supporters from the opposing sides. The background music either (1a) threatens, (1b) warns or (1c) mocks the wrestlers. The general goal of the music is to deter the opponent from being victorious. As Apter (1998) rightly puts it, the wrestlers fight with their bodies, but their supporters fight with words. Similar behaviours are found in non-African events and games (see Lee, 1985; McLeod, 2006; Vale and Fernandes, 2018).

(1) Songs from Oròyèyè festival Wrestling matches (source Apter, 1998)

- |    |         |  |   |
|----|---------|--|---|
| a  | Threat  | bírí gbe.<br>òyì gbe...                        | ‘be carried by the wind’<br>‘be dizzy’                          |
| b. | Warning | omodé é gun’gi ògèdè,<br>a yò báará, aṣe wì... | ‘a child that climbs the banana tree,’<br>‘will slide down’     |
| b. | Mockery | Adé ró di,<br>bí oní kọyín                     | ‘the crown has thumped down,’<br>‘like a bunch of palm kernels’ |

Musical 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 masqueraders. When the phrases in (2) are played on the drum (or spoken by the spectators), the masqueraders 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).

(2) Yorùbá drum phrases (Fámúle, 2018)

- |    |  |   |
|----|--|---|
| a  | o ò le ṣe bí baba ṣe ní ṣe                             | ‘you are not as competent as your father’                   |
| b. | b’óbá ṣe pé’mi nì’wọ ni<br>n̄ bá fá pá jó, fá pá jó... | ‘If I were you’<br>‘I will dance with my hands unceasingly’ |

Thematically, the linguistic interpretations of the video game music are comparable to those of the background music in traditional Yorùbá games and entertainments. For the Yorùbá gamers, the Overworld and Underworld themes are death threats to the player-controlled character or more specifically the player. While the Death theme mocks the death of the player-controlled character, the Flagpole theme is a celebratory theme for the

victory of the player-controlled character. Based on the thematic similarities, we can say that the gamers consumed the video game music using their background knowledge about the functions of background music in Yorùbá recreational activities. 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 organizations...about what people should do” (Hasegawa et al. 2007: 180).

That the Yoruba gamers utilised their cultural knowledge about game music in their interpretation of the video game music is in line with the enculturation account of musical interpretation, which suggests that musical meaning is exclusively determined by cultural convention and social background (Keil and Feld, 1994; Walker, 1996; Gregory and Varney, 1996; Balkwill and Thompson, 1999). 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. However, Gregory and Varney (1996) suggests 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).

In addition to social expectation, the physical properties of music such as pitch and tempo have also been shown to prime the interpretation of music. For example, faster tempo are associated with joy and happiness for 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, these studies are based on music traditions without speech surrogates. Considering that Yorùbá has a music tradition with a speech surrogate system which mostly relies on mapping the pitch contours of speech tones to music melodies (Villepastour, 2010, 2014; Akinbo, 2019), it is important to inquire whether pitch contours of music melodies influence the interpretation of music. Specifically, does the music interpretation involve vocal imitation of non-linguistic sounds? For this purpose, we must compare the pitch contours of the music notes in Super Mario to those of their conventionalised linguistic interpretation in Yorùbá. This investigation is crucial to tones of the lexical syllables that are mapped to each music note. In order to compare the pitch contours of the theme music and the linguistic interpretation, I conducted an acoustic investigation on the pitch contours of the conventionalised interpretations and the theme music.

Before turning to the acoustic investigation, the summary of the discussion here is that Yorùbá gamers assume that the background music of the game is communicative as a result of drawing on the roles of background music at traditional Yorùbá games. In this case, the cultural background and experience of the gamers are motivations for the linguistic interpretation of the instrumental music.

#### 4. Methodology

Ten native speakers of Yorùbá were recruited in Vancouver, British Columbia for this study. The participants are 4 females and 6 males who immigrated from Nigeria and have spent at most seven years in Canada. They were all between the ages of 26 and 45. The stimuli in this study are the relevant Super Mario music and their conventionalised linguistic interpretations in Yorùbá. Because the sounds in Super Mario Bros often overlap, the relevant Super Mario theme music were extracted from YouTube videos of the instrumental music (sources Youcanplayit, 2012a,b,c; Luuul's, 2016). Just as in the game, the themes were performed on a piano by the YouTube sources.

The linguistic interpretations with and without the corresponding instrumental music of Super Mario were presented to each of the participants. When the participants were presented the linguistic interpretations without the instrumental music, they were instructed to read the linguistic interpretations in speech mode. When the linguistic interpretations were presented with the instrumental music, the participants were instructed to produce the linguistic interpretations in a cappella mode. Each stimulus was repeated three times in speech mode as well as in a cappella 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. The vowels in speech and a cappella renditions were annotated in Praat (Boersma, 2001). The music notes of the theme music were also annotated in Praat.  $F_0$  values of the annotated intervals were automatically extracted at 50% point for the Super Mario Bros themes and their linguistic interpretations. In the next section, I present the results of the production experiment.

#### 5. Pitch of Super Mario music and the corresponding Yorùbá interpretations

The  $F_0$  contours of the Super Mario theme music and their corresponding linguistic interpretation are presented in this section. We will start the discussion by looking at the Underworld theme. As shown earlier, the first-six notes of Underworld theme are repeated once (see §2). The pitch contour of the Underworld tune and that of its repetition are compared with their corresponding linguistic interpretations. For ease of identification, the initial portion of the Underworld theme is labelled *Underworld1* and its repetition is labeled *Underworld2*. The line plot of *Underworld1* music notes and the a cappella of its linguistic interpretation are presented in Figure 6. The y-axis contains the  $F_0$  values, and the x-axis contains the string of speech tones or music notes. The results from each participant are grouped by panels (i.e. p1, p3, p#).



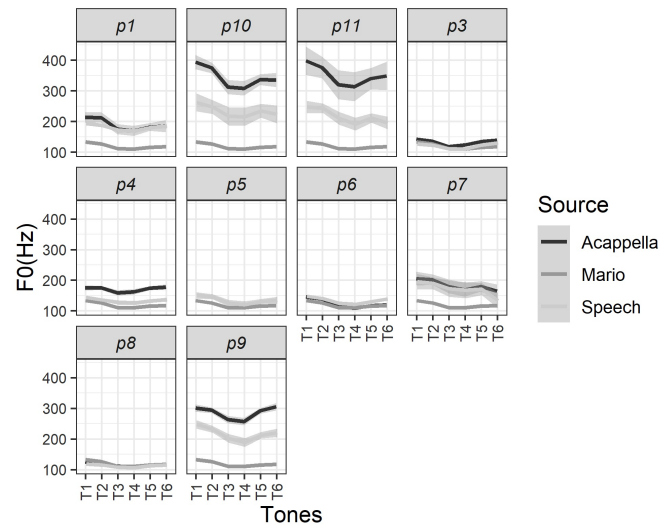


Figure 6: Pitch contours of Underworld1 in Mario and a cappella mode

The  $F_0$  contours of Underworld1 notes and those of the linguistic interpretation in an a cappella rendition are plotted in Figure 6. Apart from individual differences among the participants, the  $F_0$  contours of Underworld1 is similar to those of their linguistic interpretation in an a cappella mode for all the participants.

To understand the strength of the relationship between the  $F_0$  of Underworld1 notes and the corresponding linguistic interpretation in an a cappella mode, a correlation coefficient of the speech and music tones were calculated using *ggpubr* (Kassambara, 2018). For this purpose, the mean of  $F_0$  values for all the participants was plotted against  $F_0$  values of the music tones. The y-axis contains the  $F_0$  values for the Underworld1 notes and the x-axis contains the mean  $F_0$  values of the corresponding tones in a cappella mode. The mean values of Underworld1 notes for all the participants in the game are plotted against those of their corresponding a cappella mode. This is repeated for the other theme music and their corresponding interpretations. The result of the Underworld1 theme is presented in Figure 7. For the reasons of formatting and for having similar results, the correlation coefficient for Mario instrumentals and their interpretations in speech mode are only presented at the end of this section.

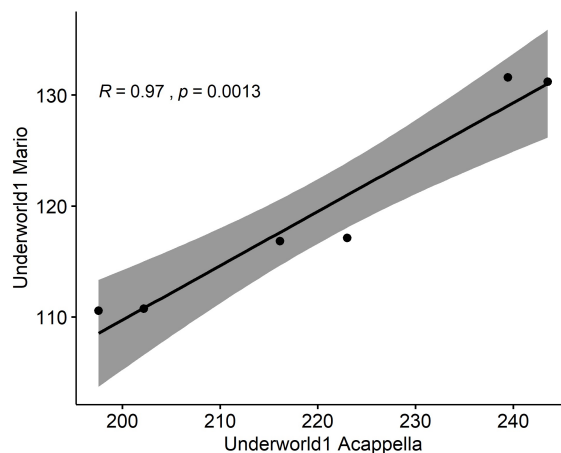


Figure 7: Mean  $F_0$  correlation between Underworld1 of Mario and a cappella mode

Figure 7 shows that there is a high degree of positive relationship ( $R=0.97$ ) between the mean  $F_0$  of Underworld1 notes and that of their corresponding linguistic interpretations in a cappella mode, and that this relationship is significant ( $p=0.0013$ ). I now turn to  $F_0$  contours of Underworld2, which is based on the repetition of the music notes in Underworld1.

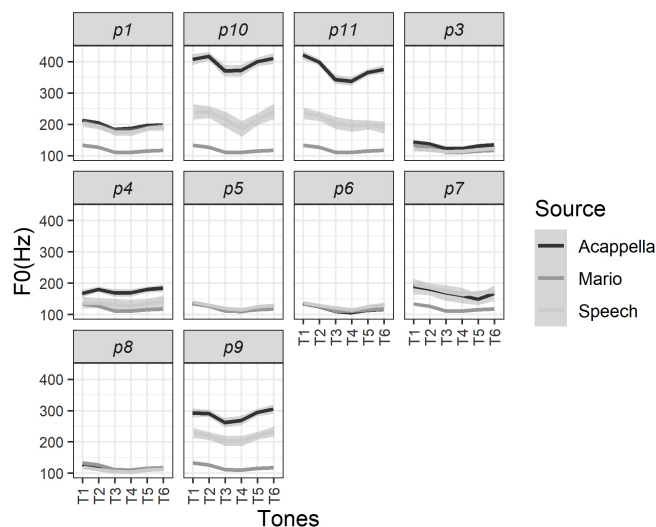


Figure 8: Pitch contours of Underworld2 in Mario and a cappella mode

Figure 8 shows that the  $F_0$  contours of Underworld2 notes and those of their linguistic interpretation in a cappella mode are similar. The degree of relationship between the  $F_0$  values of the Underworld2 notes and those of the tone in their corresponding a cappella mode are presented in Figure 9.

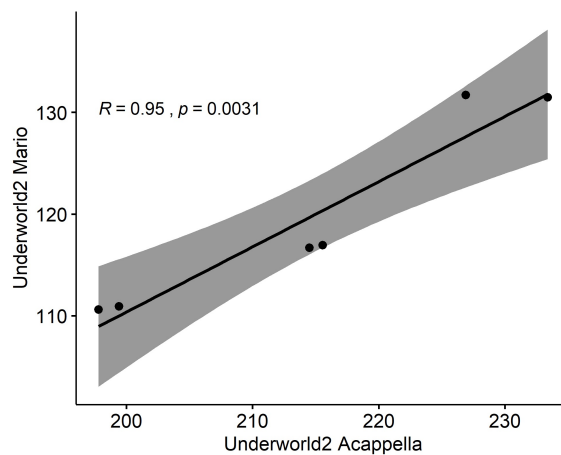


Figure 9: Mean  $F_0$  correlation between Underworld2 of Mario and a cappella mode

As shown in Figure 9, there is a high degree of positive relationship ( $R=0.95$ ,  $p=0.0031$ ) between the Underworld2 notes and their interpretation in a cappella mode.

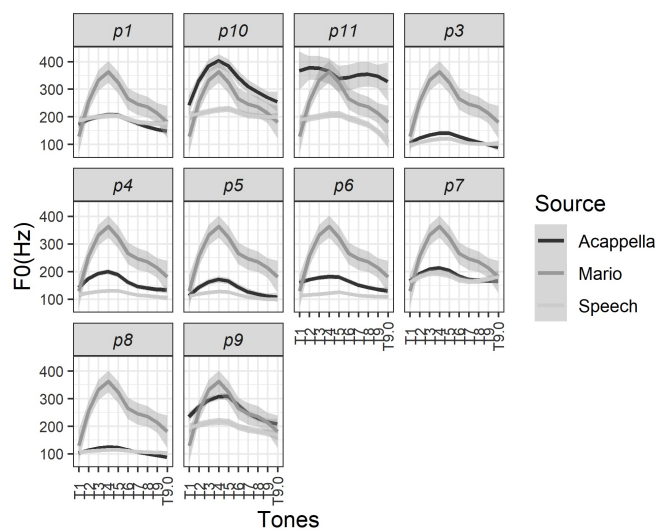


Figure 10: Pitch contours of Overworld in Mario and a cappella mode

As shown in Figure 10, the trajectory of mean  $F_0$  contours for the Overworld theme is similar to those of the linguistic interpretation in a cappella mode.

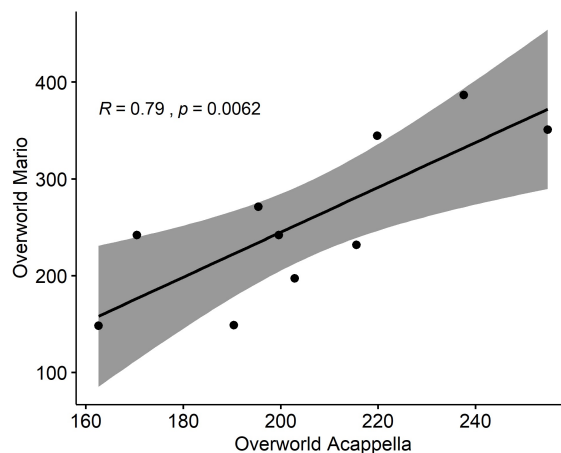


Figure 11: Mean  $F_0$  correlation between Overworld of Mario and a cappella mode

Figure 11 shows that there is a high degree of positive relationship ( $R=0.79$ ) between the mean  $F_0$  of Overworld notes and the corresponding linguistic interpretation in an a cappella mode, and that this relationship is statistically significant ( $p=0.0062$ ).

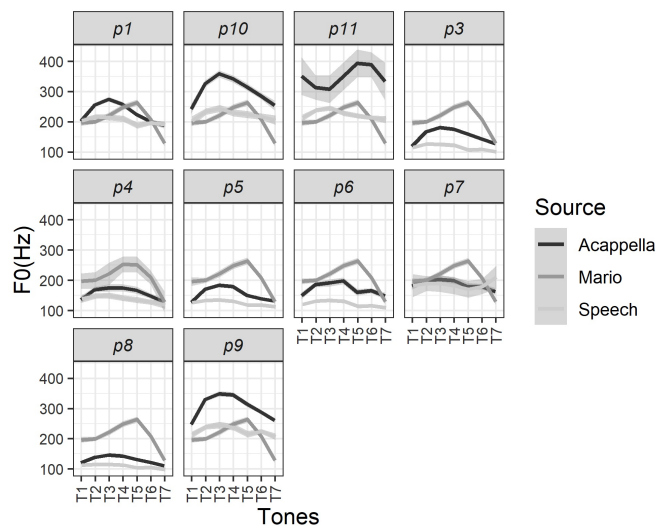


Figure 12: Pitch contours of Death theme in Mario and a acappella mode

Just as the aforementioned cases, the mean  $F_0$  contour of the Death theme is similar to those of the linguistic interpretation in an a cappella mode.

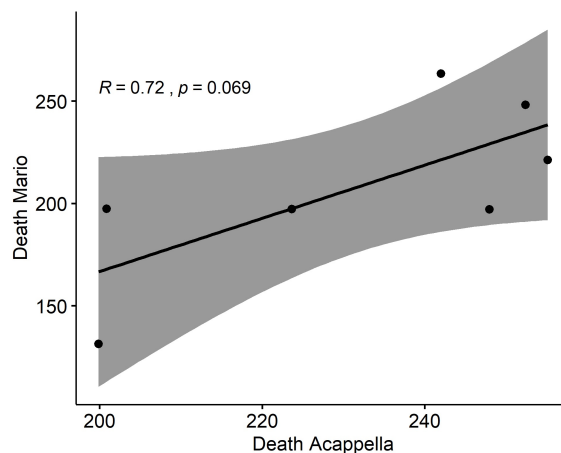


Figure 13: Mean  $F_0$  correlation between Death of Mario and an a cappella mode

As shown in Figure 13, there is a high degree of positive relationship ( $R=0.72$ ) between the mean  $F_0$  of Death theme and the corresponding linguistic interpretations in an a cappella mode, but the relationship is not statistically significant ( $p=0.069$ ).

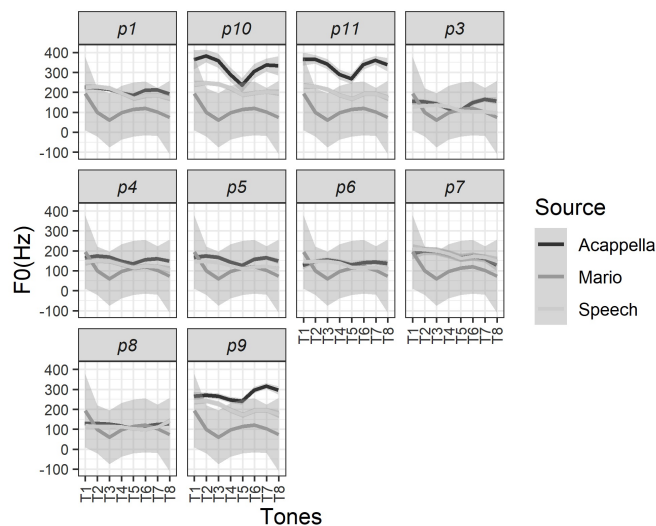


Figure 14: Pitch contours of Flagpole theme in Mario and a capella mode

Unlike the results of the previous theme music, the mean  $F_0$  contour of Flagpole theme and that of the linguistic interpretation are mainly similar at the beginning and the end. The strength of the relationship between mean  $F_0$  of Flagpole theme and that of the corresponding linguistic interpretation in a cappella mode is shown in Figure 15.

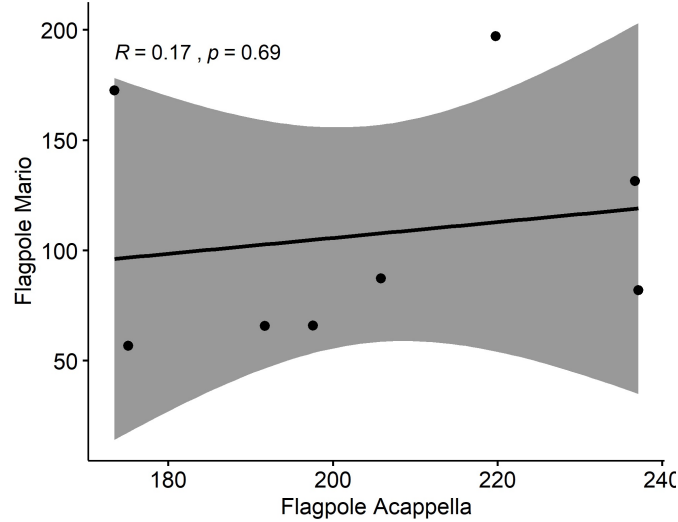


Figure 15: Mean  $F_0$  correlation between Flagpole of Mario and a cappella mode

As shown in Figure 15, there is a low degree of positive relationship between mean  $F_0$  values of the Flagpole notes and those of the corresponding linguistic interpretation in a cappella mode ( $p=0.17$ ,  $p=0.069$ ).

So far, we have only compared Mario theme music to a cappella mode of the corresponding linguistic interpretations. This is because the results are similar when we compare Mario theme music to the speech of the corresponding linguistic interpretations. The summary of all the results are presented in Table 1.

Table 1: Correlation coefficients of the music notes and their linguistic interpretations

		<i>S</i> vs <i>A</i>	<i>M</i> vs <i>S</i>	<i>M</i> vs <i>A</i>
Underworld1	<i>R</i>	0.97	0.94	0.97
	<i>p</i>	0.0011	0.0046	0.0013
Underworld2	<i>R</i>	0.95	0.93	0.95
	<i>p</i>	0.0037	0.0073	0.0031
Overworld	<i>R</i>	0.93	0.67	0.79
	<i>p</i>	0.00011	0.034	0.0062
Death	<i>R</i>	0.75	0.28	0.72
	<i>p</i>	0.052	0.54	0.069
Flagpole	<i>R</i>	0.74	0.22	0.17
	<i>p</i>	0.034	0.59	0.69

The summary of the results is as follows. There is a positive linear relationship between the pitch contours of M(ario) instrumentals and their linguistic interpretations in A(cappella) or S(peech) mode. Relative to the pitch contours of the speech renditions, the pitch contours of the a cappella renditions are closely related to the pitch contours of the

Mario instrumentals. With the exception of the Death and Flagpole tunes, the relationships between the pitch contours of all the music themes and their linguistic interpretations are statistically significant. The results also show that there is a positive relationship between the pitch contours of the interpretation in speech and a cappella modes, but the degree of similarity varies depending on the Mario instrumentals which is interpreted.

### **5.1 Linguistic interpretation of music involves acoustic mapping**

The results of the acoustic analysis show that the pitch contours of the Super Mario theme music are similar to those of their linguistic interpretations, but depending on the theme music, the strength of this relation is either strong or weak. To explain the similarity, it would be useful to refer to the talking drum rendition of Yorùbá words. As reported in Akinbo (2019), Yorùbá drummers communicate by distinctively representing the pitch contours of tones in Yorùbá with their native drums. 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. That the Yorùbá drummers musically encode the tones in Yorùbá is an example of 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).

Comparatively, in the linguistic interpretation of Super Mario music and the talking drum rendition of Yorùbá words, there is a positive relationship between the pitch contours of non-vocal music and those of their corresponding linguistic forms. Based on this similarity, we can say that the linguistic interpretation of instrumental music, just as the drumming of Yorùbá words, is based on sound imitation or phonetic iconicity.

Despite the similarity between the linguistic interpretation of the Super Mario music and the talking drum rendition of Yorùbá words, there are certain differences. Tone-based speech surrogates do not encode segmental features (McPherson, 2019), but the vocal imitation of the Super Mario music contains segmental features. If we take into account the fact that vowel and consonant types affect the pitch value of a tone (Hombert, 1977; Whalen and Levitt, 1995; Whalen et al., 1999), the positive relationship between the pitch contours of the linguistic interpretations and those of the corresponding non-vocal music being strong or weak could be an effect of segmental features on the pitch value of a tone. When we compare the phonetic properties of the linguistic interpretation to those of the Super Mario theme music, we observe that the music notes of the Super Mario theme music do not contain information about segmental properties (see Jackendoff and Lerdaahl, 2006; Jackendoff, 2009). So, this begs the question: what determines the segment or morpheme that is matched to the musical patterns of the Super Mario theme music? In the next section, we focus on this issue.

Another interesting aspect of the results involves the correlation between the pitch trajectories of the interpretations in speech and a cappella modes. Although there is a positive correlation between the speech and a cappella of the musical interpretations, the strength of the relationship varies depending on the Mario instrumentals which is musi-

cally interpreted. In fact, the correlation between the pitch contour of the speech and a cappella 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).

Before turning to the segmental properties of the vocal imitation, the discussion in this section is summarised as follows. Yorùbá gamers linguistically interpreted the Super Mario theme music by mapping the pitch contours of the music notes to Yorùbá phrases with similar pitch contours. Specifically, the gamers achieved this pitch-contour matching by modifying a suitable Yorùbá idiom or generating strings of words with similar pitch contours as the instrumentals.

## **6. Situational contexts contribute to linguistic interpretation of music**

I account for the segmental or lexical properties of the linguistic interpretation of the Super Mario theme music in this section. For this purpose, we must refer to the issue of ambiguity in drum language. In phoneme-based speech surrogates, ambiguity is a current problem, and it is attributed to the drum language relying solely on tones of the source language (Ajayi, 1990; Poss, 2005; Pulleyblank, 2009). In speech surrogate research, there is a consensus that situational contexts resolve the recurrent problem of ambiguity in drum communication (Stern, 1957; Euba, 1990; Agawu, 1995, 2001).

The ambiguity of drummed speech and its resolution are clearly illustrated in the theme of “Oníbodè Lálúṣon”, a poem written by the Yorùbá writer, Adébáyò Fálétí (Ọlátúnjí 1982: 2-5). The poem is about an event between an unattractive wealthy man called Oníbodè Lálúṣon and a drummer who prefers to only play for wealthy people. When the drummer passes by the house of Oníbodè Lálúṣon, he gifts money and palm-wine to the drummer. To acknowledge Oníbodè Lálúṣon’s generosity, the drummer plays a tune on his drum (3a). Oníbodè Lálúṣon interprets the drum tune as a praise tune (3b), but the detractors of the drummer tell Oníbodè Lálúṣon that the tune is an insult to his unattractiveness (3c). When Oníbodè Lálúṣon realises that the two interpretations of the drum tune are possible, he questions the drummer about the intended meaning of the drum tune. The drummer responds that his interpretation is of praise not insult. By taking into account the situational context of “gifting and appreciation”, Oníbodè Lálúṣon accepts the interpretation of the drummer.



## (3) Drummed speech in Oníbodè Lálúṣon

a. DRUM TUNE	b. PRAISE INTERPRETATION	c. INSULT INTERPRETATION
MMMLLL	Mo jẹun Èjìgbò <i>'I ate Ejigbo's food'</i>	E wẹnu ìmàdò <i>'Look at the mouth of the warthog'</i>
MMMLLL	Mo jẹun Ìwó <i>'I ate Iwo's food'</i>	E wẹnu Ìsín <i>'Look at the mouth of the minnows'</i>
MMMMHMLHHM	Mo jẹun Oníbodè Lálúṣon <i>I ate Oníbodè Lálúṣon's food</i>	E wẹnu Oníbodè Lálúṣon <i>'Look at Oníbodè Lálúṣon's mouth'</i>

Similar to how ambiguity is resolved in the poem, the gamers also utilised the situational contexts in the Super Mario game to determine the segments or words which are mapped to the theme music of the game. For example, the phrase “(s)he/it entered and cannot definitely come out” (4a) is mapped to the Flagpole fanfare theme which plays immediately Mario enters the castle. If we consider that the linguistic interpretation of the Flagpole theme involves mapping the sequence of tones H M H L L M H M to the music notes of the Flagpole theme, any phrase with the same tone melody could have been mapped to the music (4b-d). Mapping the phrase “(s)he/it entered and cannot definitely come out” to the Flagpole theme music suggests that the linguistic interpretation of the Flagpole theme is conditioned by the context of Mario “entering the castle and not coming out”.

## (4) Possible interpretations of the Flagpole theme

Tone: H M H L L M H M

- a. **Ó wọ'lé kòdè le jáde**      ‘he/she/it entered, cannot come out’
- b. Ó ta mí tètè lọ w'ágbo      ‘he/she/it stung me, quickly look for herbs’
- c. Déwọlé àgbàdò dára      ‘Déwọlé, corn is good’
- d. etc.

Similar to the interpretation of the Flagpole theme, the situational context in the game is also utilised in the interpretation of the Death theme. For instance, the Death theme is played when Mario dies. Mapping the phrase “it is a dangerous game that you came to play” (5a) to the music notes of the Death theme correlates with mocking the death of Mario for playing a dangerous game. The theme of danger is also echoed in the linguistic interpretations of the Overworld theme (5b) and the Underworld theme (5c).

## (5) The interpretations of other Super Mario themes

a. **Death theme**

Tone: M H H H M H M

erékéré lo wa şe ‘it is a dangerous game that you came to play’

b. **Overworld theme**

Tone: H M H L L M H M

tẹ̀bà láyẹ̀ ko wá gbọ̀bẹ̀ l’ọ̀run ‘make ẹ̀bà on earth, come to heaven for soup’

c. **Underworld theme**

Tone: H H M M H H

nǐbí lo ma kú sí ‘you are going to die here...’

bóyá lo ma dé ’lé ‘(I) don’t think you will get home’

The Overworld theme is the main theme of the game, and it is the theme which is heard at the beginning of level one. The interpretation of the Overworld theme, which is “make ẹ̀bà on earth and come to heaven for soup” (5b), 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 theme as a warning or threat which is meant 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 theme plays immediately at the beginning of the second level which is set in the Underworld. Interpreting the Underworld theme as “you are going to die here” is a threat to the player who escaped death at level one.

As mentioned in §2, the Underworld theme involves the repetition of the first six notes. Although the linguistic interpretation of the music melody and its repetition which is “I don’t think you will get home” are semantically different, they have the same communicative functions because both interpretations are death threats to the player. This suggests that two linguistic interpretations are possible for the same tone pattern inasmuch as each interpretation matches the situational context of music performance.

The summary of the discussion in this section is that the gamers utilised the situational contexts and visual imagery of the game in their interpretation of the Super Mario theme music. More specifically, the situational contexts of the musical performance assigns meaning to the vocal imitation of the instrumental music.

## 7. Formal account of musical meaning

The goal of this section is to present a formal account of how Yorùbá gamers interpreted the Super Mario theme music. 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; Neumeier, 2015). In this

section, we will briefly discuss how these two approaches can account for the interpretation of Super Mario instrumentals by Yorùbá gamers.

First, we start with the semantic approach. In the semantic approach to musical meaning, musical features, such as pitch, tempo and rhythm, are capable of expressing affective and non-affective meanings. For instance, research on adult listeners has shown that faster tempo and high pitch prime affecting communication which are related to joy, happiness and pleasure and that slow tempo and low pitch prime the affective meaning of sadness (Scherer and Oshinsky, 1977; Kellaris and Kent, 1993; Gagnon and Peretz, 2003; Webster and Weir, 2005; Steinbeis and Koelsch, 2011; Eerola et al., 2013). Similar results are found in studies that focus on for children music listeners (Dalla Bella et al., 2001; Mote, 2011).

The study of musical meaning has also been extended to non-affective meanings or concepts. To determine whether unfamiliar instrumental music could suggest semantic concepts, Koelsch et al. (2004) investigate semantic priming of words using unfamiliar instrumental music excerpts and sentences. Some musical stimuli in the study are iconic (resembling bird sound, sigh, etc), while other musical stimuli have stereotypical forms or style that are associated with particular words (e.g devotion and caravan). The music-word pairing for semantic relatedness was based on ratings from a separate behavioural experiment. For semantic priming with sentences, a sentence such as “*She sings a song facilitates the processing of a semantically related word such as music, not sock*” (Koelsch et al. 2004: 302). In the experiments, the participants listened to a musical excerpt or a sentence, then they were visually presented a semantically related or unrelated word. After the visual presentation, the participants were asked whether the prime and the target were meaningfully related or not by pressing one of two buttons. Koelsch et al. (2004) utilised a neural signature of linguistic semantic processing known as the N400. The results of the study show that the participants categorised 80% of the target words correctly. The N400 priming effect in the music domain did not differ from that of the language domain. Following from this, Koelsch et al. (2004) argue that music and language have ability to activate semantic concepts. In general, the semantic approach predicts that the pitch values of a music note should have a specific meaning.

There is some 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 Hachohen 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 (e.g. happiness and sadness). If music really elicit a semantic processing, the music notes in the Underworld theme 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 Koelsch et al. (2004) 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.

As an alternative to the semantic approach, 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.

- (6) 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 themes in musical discourse, Patel (2008) suggests that investigating coherence in musical discourse could offer more valuable insights on musical meaning. The pragmatic approach predicts that the internal structure of a musical piece, such as the number of musical notes, should have an effect on musical meaning. The results of the experiments in Poss (2012) can be considered an argument in support of the pragmatic approach in Patel (2008). For example, Poss (2012) conducted an experiment on the cognition of *raj nplaim* music, which is a musical speech surrogate of Hmong people. The results of the experiment shows that the participants were able to match Hmong words to the musical melodies 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 pragmatic approach of musical meaning in Patel also has a similar limitation as the semantic approach because it suggests that musical meaning is solely dependent on the internal structure of music. Considering that the Yorùbá gamers used the context of musical performance in addition to the pitch contours the instrumental music, the prediction of the pragmatic account is also not compatible with the interpretation of Mario instrumentals by Yorùbá.

In other approaches which can also be classified as a pragmatic 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; Neumeyer, 2015). In his work on music communication, Feld (1984) argues that musical structures exist in social constructions and they have meaning through social interpretations. To account for musical meaning, Feld (1984) proposes a music communication model which is schematically represented in Figure 16. In the model, music communication process involves two dynamically linked components, namely 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: 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*.

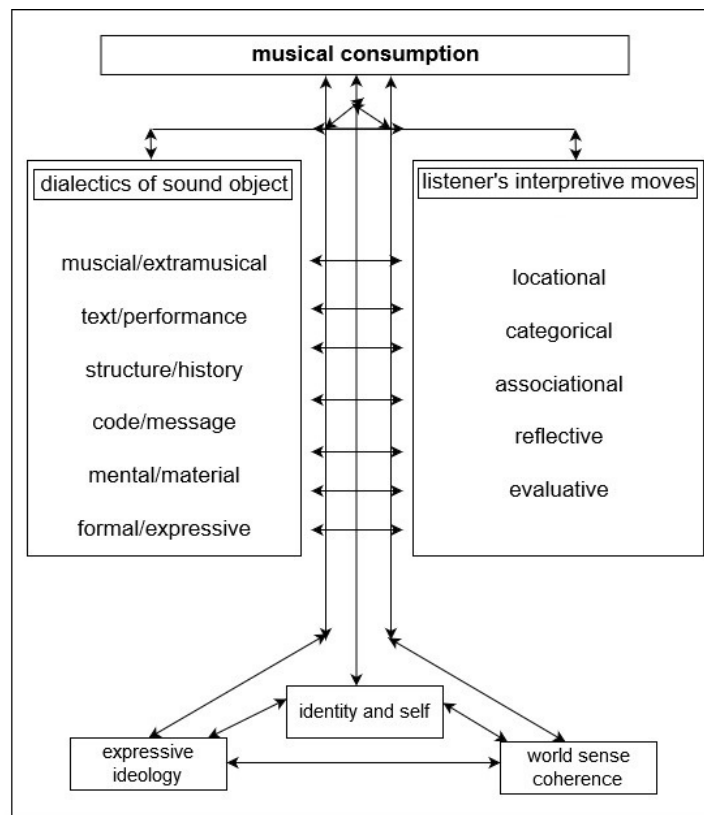


Figure 16: Interpretive moves (Feld 1984: 9)

The interpretive moves proposed in Feld (1984) can account for the interpretation of the Super Mario themes. For this purpose, we refer to the discussion in the previous section. The discussion suggests that the Yorùbá gamers used their cultural experience, pitch imitation and situational contexts of the game in their interpretation of the Super Mario instrumentals. Drawing on their native cultural experience in the interpretation of the Super Mario theme music can be considered a *reflective interpretive move*. That the gamers imitated the pitch contours of the instrumental music can be accounted for with 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 theme music of the game is 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 can be considered an effect of the associational element.

In sum, the music communicative model in Feld (1984), which can be considered a pragmatic approach to musical meaning, accounts for the interpretation of the Super Mario themes by Yorùbá gamers.

## 8. Summary, Discussion and conclusion

This study has investigated the linguistic interpretation of Super Mario musical themes by Yorùbá gamers. The results of the study suggest that the instrumentals 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 instrumentals. To determine the tone of the syllables, the second move involves mapping the pitch contours of the music melodies to Yorùbá phrases with similar pitch contours. However, the degree of similarity between the pitch contours of the music melodies and their linguistic interpretations varies depending on the music theme. 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 related to the visual events in the game and the cultural background of the listeners (i.e. the Yorùbá gamers).

That Yoruba gamers mapped the pitch contours of the music melodies to the vocal forms is in line with previous studies on vocal imitation of instrumental music (Lemaitre et al., 2016; Mantell and Pfordresher, 2013; Mercado III et al., 2014; Wisniewski et al., 2013). As reported in studies on Yorùbá talking drums, the communicative capability of the speech surrogates is based on mapping the pitch contours in speech forms to musical melodies (Villepastour, 2010, 2014; Akinbo, 2019). This suggests that the musical speech surrogates and the linguistic interpretation of music involve pitch imitation in Yorùbá.

The most important contribution of this work is the role of contextual information in the interpretation of unfamiliar instrumental music. That the Yorùbá gamers used contextual information in the interpretation of the instrumental music is an argument in support of the pragmatic approach to musical meaning (Feld, 1984; Patel, 2008; García et al., 2014; Neumeyer, 2015). While most of the studies on musical meaning or the linguistic interpretation of music are based on experimental studies, the linguistic interpretation of the Super Mario music by Yorùbá gamers presents natural evidence in support of the linguistic interpretation of instrumental music.

Given that the linguistic interpretations 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. 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 video game could be promising. Such studies will not only contribute to the fields of linguistics, psychology and music, they will contribute to the emerging field of video game music, where there is a consensus that background music of a video game 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; 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 functional, linked to dance and externally motivated by social and musical contexts (Agawu, 2006, 2001). Future research would investigate whether functional and non-functional music are interpreted differently. It would be interesting to investigate whether there is a meaningful distinction between interpreting music on the fly in the moment and a conventionalized pattern as is the case here.

It bears mentioning that only some portions of each music theme is linguistically interpreted. For future research, it would be interesting to investigate whether this has to do with the rhythmic properties of the portions which are not linguistically interpreted. If this is the case, we could investigate whether this has to do with the rhythm of these specific portions of the music themes in relation to the rhythm of speech (see Patel and Daniele, 2003; Patel, 2006, for comparative studies on rhythm in speech and music).

To conclude, Yorùbá gamers linguistically interpreted instrumental music by vocally imitating the pitch contours of musical melodies and mapping a lexical syllable to each musical note. The assignment of morpheme to the vocal imitation is through the use of the situation contexts surrounding the musical performance and the cultural background of the listener. Considering that many studies on African oral poetry do not employ video game, I recommend the use of video-game data for studies on music-language connection.

### References

- Adedeji, Joel A. 1972. The Origin and form of the Yoruba masque theatre. *Cahiers d'Études Africaines* 12(46): 254–276.
- Agawu, V Kofi. 1995. *African rhythm: A Northern Ewe perspective*. Cambridge: Cambridge University Press.
- Agawu, V Kofi. 2001. African music as text. *Research in African Literatures* 32(2): 8–16.
- Agawu, V Kofi. 2006. Structural analysis or cultural analysis? competing perspectives on the “standard pattern” of West African rhythm. *Journal of the American Musicological Society* 59(1): 1–46.
- Agawu, V Kofi. 2016. *The African imagination in music*. Oxford: Oxford University Press.
- Ajayi, Bade. 1990. Yoruba drum language. A Problem of interpretation. *Nigeria Magazine* 58: 29–37.
- Akinbo, Samuel. 2019. Representation of Yorùbá tones by a talking drum: An acoustic analysis. *LLA* : 11–23.
- Akinlabi, Akinbiyi. 1985. *Tonal underspecification and Yoruba tone*. Phd thesis, University of Ibadan.

- Apter, Andrew. 1998. Discourse and its disclosures: Yoruba women and the sanctity of abuse. *Africa* 68(1): 68–97.
- Awóbùlúyì, Oládélé. 1978. *Essentials of Yorùbá grammar*. Ibadan: Oxford University Press Nigeria.
- 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 and Jamie Russell. 2015. Expectations and speech intelligibility. *The Journal of the Acoustical Society of America* 137(5): 2823–2833.
- Balkwill, Laura-Lee and William Forde Thompson. 1999. A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception* 17(1): 43–64.
- Bamgboṣe, Ayo. 1967. *A Short Yorùbá grammar*. Ibadan: Heinemann.
- Beier, Ulli. 1954. The Talking drums of the Yoruba. *African Music* 1: 29–31.
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9–10): 341–345.
- Bradley, David. 1979. Speech through music: the Sino-Tibetan gourd reed-organ. *Bulletin of the School of Oriental and African Studies, University of London* : 535–540.
- Campbell, Bolaji. 2015. Eegun ogun: War masquerades in Ibadan in the era of modernization. *African Arts* 48(1): 42–53.
- Carrington, John F. 1971. The Talking drums of Africa. *Scientific American* 225(6): 90–95.
- Collins, Karen. 2009. An Introduction to procedural music in video games. *Contemporary Music Review* 28(1): 5–15.
- Dalla Bella, Simone, Isabelle Peretz, Luc Rousseau, and Nathalie Gosselin. 2001. A developmental study of the affective value of tempo and mode in music. *Cognition* 80(3): B1–B10.
- Devos, Thierry and Mahzarin R Banaji. 2005. American= white? *Journal of personality and social psychology* 88(3): 447.
- Olátúnjí, Olátúndé. 1982. *Ewì Adébayò Fálétí*. Ibadan: Heinemann Educational Books.
- Eerola, Tuomas, Anders Friberg, and Roberto Bresin. 2013. Emotional expression in music: contribution, linearity, and additivity of primary musical cues. *Frontiers in psychology* 4: 487.
- Ekwueme, Lazarus N. 1974. Linguistic determinants of some Igbo musical properties. *Journal of African Studies* 1(3): 335–353.
- Euba, Akin. 1990. *Yorùbá drumming: The Dùndún tradition*. Bayreuth: Bayreuth African Studies.
- Fámúle, Oláwólé. 2018. 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. *Yearbook for Traditional Music* 16: 1–18.
- Fu, Jiulin Zhang Xiaoqing. 2015. The influence of background music of video games on immersion. *Journal of Psychology & Psychotherapy* 5(4).
- Gagnon, Lise and Isabelle Peretz. 2003. Mode and tempo relative contributions to “happy-sad” judgements in equitone melodies. *Cognition and Emotion* 17(1): 25–40.
- García, Ricardo R, Francisco Zamorano, and Francisco Aboitiz. 2014. From imitation to meaning: circuit plasticity and the acquisition of a conventionalized semantics. *Frontiers in Human Neuroscience* 8: 605.
- Green, Thomas. 2005. Dambe: Traditional Nigerian boxing. *Journal of Combative Sport* .



- Gregory, Andrew H and Nicholas Varney. 1996. Cross-cultural comparisons in the affective response to music. *Psychology of Music* 24(1): 47–52.
- Grice, Herbert P. 1975. Logic and conversation 3: Speech acts. In *Speech acts*, ed. Peter Cole and Jerry L Morgan. New York: Academic Press, 41–58.
- Grimshaw, Mark. 2008. Sound and immersion in the first-person shooter. *International Journal of Intelligent Games and Simulation* 5(1): 119–124.
- Hacohen, Ruth and Naphtali Wagner. 1997. The Communicative force of Wagner’s leitmotifs: Complementary relationships between their connotations and denotations. *Music Perception* 14(4): 445–475.
- Hasegawa, Koichi, Chika Shinohara, and Jeffrey P Broadbent. 2007. The effects of ‘social expectation’ on the development of civil society in Japan. *Journal of Civil Society* 3(2): 179–203.
- Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3(1): 67–90.
- Hombert, Jean-Marie. 1977. Consonant types, vowel height and tone in Yoruba. *Studies in African Linguistics* 33: 173–190.
- Hughes, David W. 2000. No nonsense: The logic and power of acoustic-iconic mnemonic systems. *British Journal of Ethnomusicology* 9(2): 93–120.
- Jackendoff, Ray. 2009. Parallels and nonparallels between language and music. *Music perception* 26(3): 195–204.
- Jackendoff, Ray and Fred Lerdahl. 2006. The capacity for music: What is it, and what’s special about it? *Cognition* 100(1): 33–72.
- Kang, Okim and Donald L Rubin. 2009. Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology* 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 and Steven Feld. 1994. *Music grooves: Essays and dialogues*. Chicago: University of Chicago Press.
- Kellaris, James J and Robert J Kent. 1993. An exploratory investigation of responses elicited by music varying in tempo, tonality, and texture. *Journal of consumer psychology* 2(4): 381–401.
- Koelsch, Stefan. 2005. Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology* 15: 207–212.
- Koelsch, Stefan, Elisabeth Kasper, Daniela Sammler, Katrin Schulze, Thomas Gunter, and Angela D Friederici. 2004. Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience* 7(3): 302–307.
- Laroche, Guillaume. 2012. *Analyzing musical Mario-media: Variations in the music of Super Mario video games*. Dissertation.
- Lee, Martin J. 1985. From rivalry to hostility among sports fans. *Quest* 37(1): 38–49.
- Leech, Geoffrey N. 1983. *Principles of pragmatics*. London: Longman.
- Lemaitre, Guillaume, Olivier Houix, Frédéric Voisin, Nicolas Misdariis, and Patrick Susini. 2016. Vocal imitations of non-vocal sounds. *PLOS One* 11(12): e0168167.
- Lerner, Neil. 2014. Mario’s dynamic leaps: Musical innovations (and the specter of early cinema) in Donkey Kong and Super Mario Bros. In *Music in video games*, ed. Kevin J Donnelly, William Gibbons, and Neil Lerner. New York: Routledge, 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 and Matt Barton. 2012. *Vintage games: an insider look at the history of Grand Theft Auto, Super Mario, and the most influential games of all time*. Oxford: Elsevier Focal Press.
- Luuul's. 2016. Super Mario Bros – Underground theme piano tutorial. URL <https://www.youtube.com/watch?v=0iA1Tae3poc>.
- Mantell, James T and Peter Q Pfordresher. 2013. Vocal imitation of song and speech. *Cognition* 127(2): 177–202.
- McGowan, Kevin B. 2015. Social expectation improves speech perception in noise. *Language and Speech* 58(4): 502–521.
- McLeod, Ken. 2006. “We are the champions”: masculinities, sports and popular music. *Popular Music and Society* 29(5): 531–547.
- McPherson, Laura. 2018. The Talking balafon of the Sambla: Grammatical principles and documentary implications. *Anthropological Linguistics* 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 and Kevin M Ryan. 2018. Tone-tune association in tommo so (dogon) folk songs. *Language* 94(1): 119–156.
- Mercado III, Eduardo, James T Mantell, and Peter Q Pfordresher. 2014. Imitating sounds: A cognitive approach to understanding vocal imitation. *Comparative Cognition & Behavior Reviews* 9.
- 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.
- Nacke, Lennart E, Mark N Grimshaw, and Craig A Lindley. 2010. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers* 22(5): 336–343.
- Neubauer, John. 1986. *The emancipation of music from language: departure from mimesis in eighteenth-century aesthetics*. New Haven, CT: Yale University Press.
- Neumeyer, David P. 2015. *Meaning and interpretation of music in cinema*. Bloomington & Indianapolis: Indiana University Press.
- Nintendo, Entertainment System. 1985. *Super Mario Bros.: Instructional booklet*. Video game, Nintendo Co., Ltd.
- Omojola, Bode. 2011. Òṣogbo: Power, song and performance in a Yoruba festival. *Ethnomusicology Forum* 20(1): 79–106.
- Patel, Aniruddh D. 2006. Musical rhythm, linguistic rhythm, and human evolution. *Music Perception* 24(1): 99–104.
- Patel, Aniruddh D. 2008. *Music, language, and the brain*. New York: Oxford University Press.
- Patel, Aniruddh D and Joseph R Daniele. 2003. An empirical comparison of rhythm in language and music. *Cognition* 87(1): B35–B45.
- Poss, Nicholas Frederick. 2005. *The Communication of verbal content on the Hmong Raj: an ethnographic analysis of performance practice*. MA Thesis, The Ohio State University.
- Poss, Nicholas Frederick. 2012. *Hmong music and language cognition: An interdisciplinary investigation*. Phd thesis, The Ohio State University.
- Pulleyblank, Douglas. 1986. *Tone in lexical phonology*. Dordrecht, Holland: D. Reidel Publishing Company.

- Pulleyblank, Douglas. 2009. Yoruba. In *The World's major languages*, ed. Bernard Comrie. Taylor & Francis, 866–882.
- Rubin, Donald L. 1992. Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education* 33(4): 511–531.
- Sanders, Timothy and Paul Cairns. 2010. Time perception, immersion and music in videogames. *Proceedings of HCI 2010 24* : 160–167.
- Schartmann, Andrew. 2015. *Koji Kondo's Super Mario Bros. soundtrack*. New York: Bloomsbury Publishing.
- 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*. PhD Thesis, University of British Columbia.
- Scherer, Klaus R and James S Oshinsky. 1977. Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion* 1(4): 331–346.
- Seifart, Frank, Julien Meyer, Sven Grawunder, and Laure Dentel. 2018. Reducing language to rhythm: Amazonian Bora drummed language exploits speech rhythm for long-distance communication. *Royal Society Open Science* 5(4): 170354.
- Steinbeis, Nikolaus and Stefan Koelsch. 2011. Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience* 23(3): 604–621.
- Stern, Theodore. 1957. Drum and whistle 'languages': an analysis of speech surrogates. *American Anthropologist* (59): 487–506.
- Stuart, Keith. 2010. Super Mario Bros: 25 Mario facts for the 25th anniversary. URL <https://amp.theguardian.com/technology/gamesblog/2010/sep/13/games-gameculture>.
- Vale, Leonor and Teresa Fernandes. 2018. Social media and sports: driving fan engagement with football clubs on Facebook. *Journal of Strategic Marketing* 26(1): 37–55.
- Villepastour, Amanda. 2010. *Ancient text messages of the Yorùbá bàtá drum: Cracking the code*. Surrey, England: Ashgate Publishing Limited.
- Villepastour, Amanda. 2014. Talking tones and singing speech among the Yorùbá of South-west Nigeria. In *Jahrbuch des phonogrammarchivs der österreichischen akademie der wissenschaften*, ed. Gerda Lechleitner and Christian Liebl, vol. 44. Göttingen: Cuvillier, 29–47.
- Walker, Robert. 1996. Open peer commentary: can we understand the music of another culture? *Psychology of Music* 24(2): 103–114.
- Waterman, Christopher A. 2000. Yoruba popular music. In *The Garland handbook of African music*, ed. Ruth M Stone. Oxford: Routledge, 169–185.
- Webster, Gregory D and Catherine G Weir. 2005. Emotional responses to music: Interactive effects of mode, texture, and tempo. *Motivation and Emotion* 29(1): 19–39.
- Weir, Michele. 2015. The Scat singing dialect: An introduction to vocal improvisation. *The Choral Journal* 55(11): 28.
- Whalen, Doug H, Bryan Gick, Masanobu Kumada, and Kiyoshi Honda. 1999. Cricothyroid activity in high and low vowels: Exploring the automaticity of intrinsic f0. *Journal of Phonetics* 27(2): 125–142.
- Whalen, Douglas H and Andrea G Levitt. 1995. The universality of intrinsic F0 of vowels. *Journal of phonetics* 23(3): 349–366.
- Winter, Yoad. 2014. On the grammar of a Senegalese drum language. *Language* 90(3): 644–668.
- Wisniewski, Matthew G, James T Mantell, and Peter Q Pfordresher. 2013. Transfer effects in the

- vocal imitation of speech and song. *Psychomusicology: Music, Mind, and Brain* 23(2): 82.
- Yi, Han-Gyol, Jasmine EB Phelps, Rajka Smiljanic, and Bharath Chandrasekaran. 2013. Reduced efficiency of audiovisual integration for nonnative speech. *The Journal of the Acoustical Society of America* 134(5): EL387–EL393.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.
- Youcanplayit. 2012a. Super Mario Bros – Flagpole [easy piano tutorial]. URL [https://www.youtube.com/watch?v=ist0vB\\_idcs](https://www.youtube.com/watch?v=ist0vB_idcs).
- Youcanplayit. 2012b. Super Mario Bros–death–easy piano tutorial. URL [https://www.youtube.com/watch?v=c08xgXcm\\_hg](https://www.youtube.com/watch?v=c08xgXcm_hg).
- Youcanplayit. 2012c. Super Mario Bros–piano tutorial–easy. URL [https://www.youtube.com/watch?v=eQ\\_VLqJC8UY](https://www.youtube.com/watch?v=eQ_VLqJC8UY).
- Zehnder, Sean M and Scott D Lipscomb. 2006. The Role of music in video games. In *Playing video games – motives, responses, and consequences*, ed. Peter Vorderer and Jennings Bryant. Mahwah, New Jersey: Lawrence Erlbaum Associates, 241–258.