



THE DÙNDÚN DRUM HELPS US UNDERSTAND HOW WE PROCESS SPEECH AND MUSIC

Lauren Fink^{1,2*}, Cecilia Durojaye^{1,3}, Tina Roeske¹, Melanie Wald-Fuhrmann^{1,2} and Pauline Larrouy-Maestri^{2,4}

¹Department of Music, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

²Center for Language, Music, and Emotion (CLaME), Max Planck-New York University (NYU), Frankfurt am Main, Germany

³Department of Psychology, Arizona State University, Tempe, AZ, United States

⁴Neuroscience Department, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

YOUNG REVIEWERS:



ETHAN

AGE: 12



JANE

AGE: 11



JAIDEN

AGE: 15

Every day, you hear many sounds in your environment, like speech, music, animal calls, or passing cars. How do you tease apart these unique categories of sounds? We aimed to understand more about how people distinguish speech and music by using an instrument that can both “speak” and play music: the dundún talking drum. We were interested in whether people could tell if the sound produced by the drum was speech or music. People who were familiar with the dundún were good at the task, but so were those who had never heard the dundún, suggesting that there are general characteristics of sound that define speech and music categories. We observed that music is faster, more regular, and more variable in volume than “speech.” This research helps us understand the interesting instrument that is dundún and provides insights about how humans distinguish two important types of sound: speech and music.

DÙNDÚN TALKING DRUM

An hourglass-shaped wooden drum with two heads fastened to the frame with cords.

Changing pressure on the cords while striking the drum allows production of a wide range of pitches.

tone language

Languages that use pitch (or tone) to distinguish words. Yorùbá is a tone language with high, medium, and low tones, spoken in West Africa by about 50 million people.



Figure 1

(A) The Yorùbá language. The three lines in the speech bubble represent the tone levels of Yorùbá: high, medium, and low.

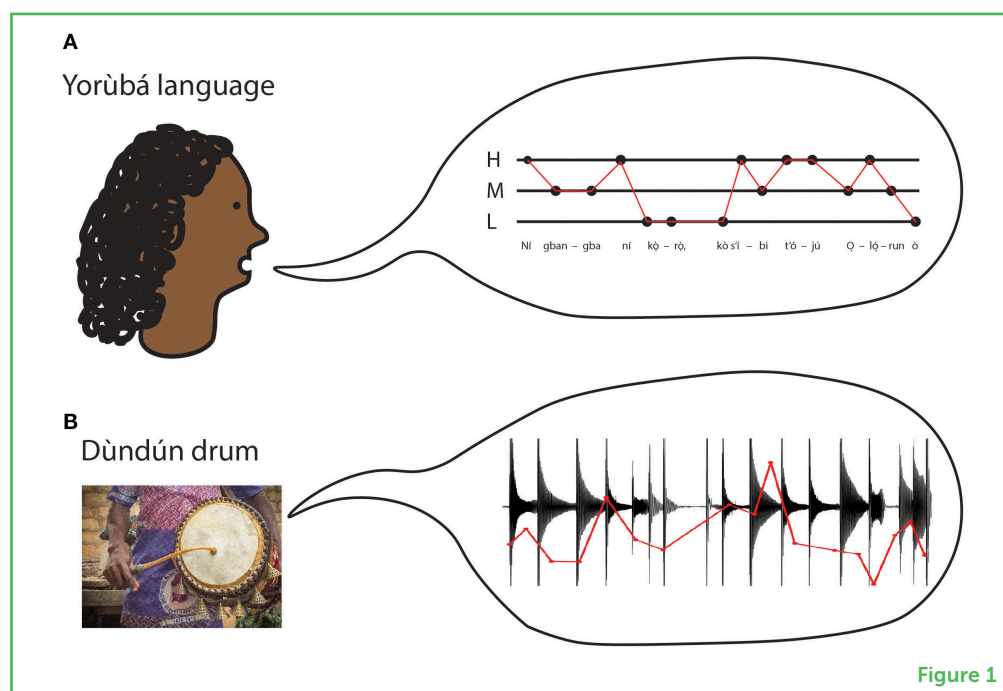
(B) The Yorùbá language can be “spoken” on the dùndún talking drum. The black line in the speech bubble shows the changes in drum intensity over time. Each burst indicates a strike of the drum. Some strikes are louder (higher) and longer in duration (black triangle shape) than others. In both speech bubbles, the red line corresponds to changes in tone. Both sound clips follow the same general changes in tone over time (Photograph Credit: Cecilia Durojaye).

A DRUM WITH ITS OWN LANGUAGE

Drums are one of the most ancient musical instruments. Early versions of drums have been discovered from as far back as 70,000 B.C.E.! Drums make loud, deep sounds that can travel long distances. Therefore, drumming has been used not only in enjoyable music-making contexts or to accompany celebrations, but also as a means to communicate information.

Here, we focus on a particular drum called the **dùndún talking drum**, which comes from southwest Nigeria, a country in Africa. In Nigeria, many people speak a language called Yorùbá. Unlike English, to speak Yorùbá, certain parts of words must be spoken at particular pitches to be properly understood. This is the principle of a **tone language**. Yorùbá is not the only tone language—Mandarin Chinese is another example. In the Yorùbá language, there are three distinct tone levels: low, medium, and high. The dùndún can imitate those tone levels so that it sounds similar to a person talking. For example, in Yorùbá you could say:  [https://lkfink.github.io/assets/audio/Yoruba_example.ogg] and on the drum you would play  [https://lkfink.github.io/assets/audio/dundun_example.ogg].

Can you hear the similarity between the two examples? Even though one is the human voice, and the other is a drum, the tones and timing of the two examples should sound similar to you. If you are having trouble hearing the changes over time, try following along with the red lines in Figure 1 while you listen.



IS MUSIC OR SPEECH COMING FROM THE DRUM?

Understanding exactly what the talking drum is saying (the content of its message) would require being familiar with the Yorùbá language. However, we were interested in whether people would be able to tell, in general, whether the drum was speaking or playing music. Do you think you could tell the difference? Try it! Below we provide four examples of dùndún performances. Can you guess whether each one is music vs. speech?¹

🔊 [https://lkfink.github.io/assets/audio/D8M_kids.ogg]

🔊 [https://lkfink.github.io/assets/audio/D7s_kids.ogg]

🔊 [https://lkfink.github.io/assets/audio/D13s_kids.ogg]

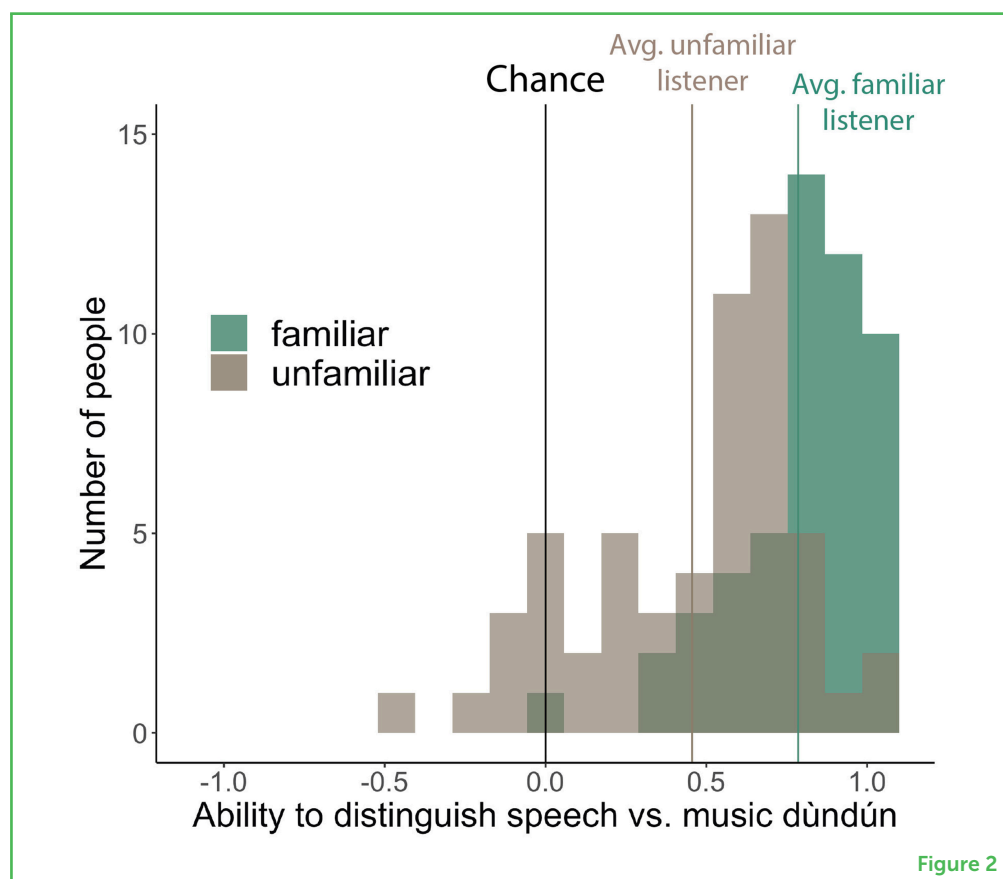
🔊 [https://lkfink.github.io/assets/audio/D6M_kids.ogg]

We asked over 100 people to do a similar task. About half of them were familiar with the dùndún or the Yorùbá language and the other half were not. All of them listened to 30 different dùndún performances and guessed whether they were speech or music. We computed a score that shows how well each participant identified the sounds as music versus speech. As you can see in Figure 2, scores can range from

¹ An answer key can be found at the end of the article or at the bottom of this webpage: <https://lkfink.github.io/pages/dundunKids>.

Figure 2

People listened to dùndún performances and classified them as speech-like or music-like. Performance was plotted on a scale from -1 to 1, with -1 meaning people guessed worse than chance, 0 being chance, and 1 representing a perfect ability to distinguish speech- vs. music-like dùndún performances. Regardless of whether they were familiar (teal) or unfamiliar (brown) with the dùndún, or the Yorùbá language, most people were able to classify dùndún performances as speech-like or music-like higher than the chance level.



–1 (every sound was misclassified) to 1 (every sound was perfectly classified); 0 corresponds to guessing randomly. As expected, we found that those familiar with the dùndún and/or those that could speak Yorùbá could tell the difference between drum speech and drum music—ten of them classified the sounds perfectly, and the others made only a few errors. Interestingly, even non-familiar listeners could tell which dùndún performances were speaking or playing music, much better than if they were just guessing.

But how did they know? What were the differences in the sounds that informed listeners' decisions about whether they were hearing speech or music? Were they using the tone (also known as **pitch**) of the notes they heard? What about the timing of the notes? Did the loudness of the notes matter? Did speech drumming vs. music drumming show specific patterns of change in any of those characteristics over time? To answer such questions, we used a computer to analyze the **acoustic** characteristics of the 30 dùndún performances.

For every note played on the dùndún, we extracted its pitch, duration, and loudness, and we analyzed general characteristics across the entire performances of speech-like vs. music-like dùndún. To do so, we used computer programs, but we also listened to and looked at the **waveforms** (a representation of a sound as changes in loudness over time), to make sure that the program did not make mistakes detecting the acoustic information. We found that, when the drum was speaking, the sound was generally louder, the pitch was lower, and the gap in time between consecutive notes, referred to as the **inter-onset interval (IOI)**, was longer. We also looked at differences between consecutive notes, for example, how much variation in loudness occurs between them. We found that, when the drum was speaking, the loudness of the sounds did not change as much as when it was playing music. Also, the timing of the speaking drum was more irregular. See Figure 3 for a summary of the measures we used and how each relate to music-like (blue) vs. speech-like (orange) performances on the drum.

We then analyzed how our participants, both familiar and unfamiliar with the instrument or the language, used those acoustic characteristics to decide whether they were hearing music or speech. Interestingly, we found that, even though both groups did well on the task, they seemed to be using the acoustic characteristics of the waveform differently. For example, people familiar with the dùndún or the Yorùbá language seemed to rely more strongly on intensity and timing when they made their judgements than did those unfamiliar with them. These results teach us that a person's familiarity with a sound source, in this case the dùndún drum or the Yorùbá language, leads them to use acoustic characteristics differently when they try to decide what they are hearing.

PITCH

The perceived highness or lowness of a sound. The **dùndún** produces a higher-pitched sound when the drum head is tighter (more pressure on cords) and vice versa.

ACOUSTIC

This word indicates that something is related to sound or the sense of hearing. Acoustic characteristics of a sound can be analyzed and quantified with computers.

WAVEFORM

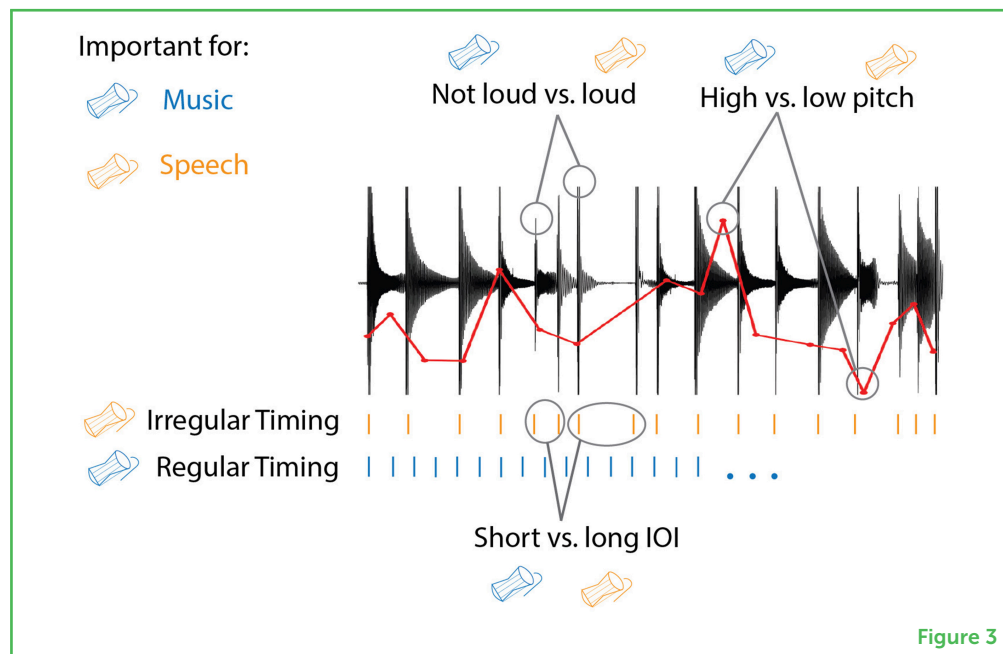
A representation of a sound as changes in loudness over time. The larger the peaks are, the more intense (louder) the sound is perceived by a listener.

INTER-ONSET-INTERVAL (IOI)

The time between the starts of sounds, measured in milliseconds. When IOIs are regular, they create a beat. Shorter IOIs are typically perceived as fast and longer IOIs as slow.

Figure 3

Waveform (black) and pitch contour (red) of a dùndún performance. Gray circles illustrate the acoustic information used to quantify loudness, pitch, and IOI. The timing ticks directly below the waveform represent the onset of each note. An IOI is long when there is a large space between two ticks. The ticks in the first row are irregular, while the second-row ticks illustrate more regular timing. The analyses of the dùndún performances showed that the drum was louder, higher in pitch, slower, and less regular when it was speaking (orange dùndún) than when it was playing music (blue dùndún).



THE DÙNDÚN CAN HELP US UNDERSTAND HUMAN PERCEPTION

Drums are not only great instruments to make people dance or to communicate a message, they can also be used to understand how humans perceive and process sounds [1]. One question scientists are currently studying concerns the relation between sounds and concepts (ideas represented by sound information). To understand this relationship, it is necessary to study various types of sounds and concepts. If you think about music and language, they both use sounds that can lead to multiple reactions, from experiencing joy, to learning new information, to recognizing something familiar. A symphony and your teacher's voice are both waveforms that enter your ears and are decoded by your brain. While the messages they transmit are obviously different, scientists are still trying to identify what is the same and what is different when our brains process music and speech signals [2, 3] and what affects our perception of them [4].

To answer this question, we used an instrument that intertwines language and music: the dùndún! While we could have used the human voice (which can speak or sing), we chose not to because people become familiar with human voices from an early age and develop specific strategies to decide whether the voice they are hearing is speaking or singing. By using the dùndún talking drum, we could identify acoustic features that shape our mental categories of music vs. speech. For example, speech-like dùndún performances are generally slower and show more variability in note onset over time, which is a dynamic (changing over time) characteristic. We think that an increased focus on dynamic features is necessary to more fully understand how sounds come to be associated with music vs. speech.

We also discovered that, while there may be general properties of sounds that are associated with music vs. speech, a person's previous experience shapes the way they use those acoustic features. In other words, there is not a simple, straight-forward relationship between acoustic cues and the mental categories of sounds, so we still need to find out what the mediating factors are.

Thus, while the perception of sounds and their association with meanings may seem simple because you do it every day, it is actually quite a complex phenomenon that relies on many different processes, from the movement of sounds from your ear to your brain, to the way you use your past experience to understand what you are currently hearing. There is still much work to do before we fully understand these elements, but the good news is that we can answer such big scientific questions using something as fun as drums!

ORIGINAL SOURCE ARTICLE

Durojaye, C., *Fink, L., Wald-Fuhrmann, M., Roeske, T., and Larrouy-Maestri, P. 2021. Perception of Nigerian talking drum performances as speech-like vs. music-like: the role of familiarity and acoustic cues. *Front. Psychol.* 12:652673. doi: 10.3389/fpsyg.2021.652673

AUDIO FILES

All audio files linked in this article can be accessed online at: <https://lkfink.github.io/pages/dundunKids>.

ANSWER KEY

Music, Speech, Speech, Music.

REFERENCES

1. Henning, D., Sabic, E., and Hout, M. 2018. Hear and there: sounds from everywhere! *Front. Young Minds* 6:63. doi: 10.3389/frym.2018.00063
2. Desai, M., Sorrells, R., Leonard, M., Chang, E., and Hamilton, L. 2020. Brain stimulation can help us understand music and language. *Front. Young Minds* 8:16. doi: 10.3389/frym.2020.00016
3. Patel, A. D. 2010. *Music, Language, and the Brain*. New York, NY: Oxford University Press.
4. Deutsch, D., Henthorn, T., and Lapidis, R. 2011. Illusory transformation from speech to song. *J. Acoust. Soc. Amer.* 129:2245–52. doi: 10.1121/1.3562174

SUBMITTED: 08 August 2021; **ACCEPTED:** 15 September 2022;

PUBLISHED ONLINE: 26 October 2022.

EDITOR: Shekher Mohan, Liberty University, United States

SCIENCE MENTORS: Christine Kurlawalla-Martinez and Nathan A. Jorgensen

CITATION: Fink L, Durojaye C, Roeske T, Wald-Fuhrmann M and Larrouy-Maestri P (2022) The Dùndún Drum Helps Us Understand How We Process Speech and Music. *Front. Young Minds* 10:755390. doi: 10.3389/frym.2022.755390

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

COPYRIGHT © 2022 Fink, Durojaye, Roeske, Wald-Fuhrmann and Larrouy-Maestri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

YOUNG REVIEWERS



ETHAN, AGE: 12

I am fascinated by all topics in STEM, especially 3D-printing, robotics, and astronomy. My hobbies include Legos, playing card games, learning magic tricks, and watching the Office on Netflix.



JANE, AGE: 11

I love science because it helps me understand things and how they work. I want to learn more about animals because I think they are wonderful creatures and I really want to protect them. I love playing with cats, reading a lot (mostly fantasy books), and playing the ukulele.



JAIDEN, AGE: 15

I have been interested in science since the age of 7. My favorite journal to read is *Scientific American*. My academic interests include chemistry, economics, and entrepreneurship. My hobbies include equestrian show jumping, playing board & card games, puzzles & riddles.

AUTHORS



LAUREN FINK

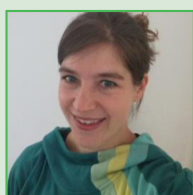
From a young age, I played the drums. I went to music conservatory to become a professional musician and became fascinated with how music affects the mind and body. Therefore, after conservatory, I pursued a Ph.D. in neuroscience. I aim to understand how music affects the nervous system, and, conversely, how the nervous system shapes what we perceive and pay attention to. My research has

taken me from Ohio (USA), to England, California (USA), and Germany, where I am currently a researcher at the Max Planck Institute for Empirical Aesthetics. *lauren.fink@ae.mpg.de



CECILIA DUROJAYE

I grew up in Nigeria in a musical family and started singing in a church choir at age eight. One year later, I began to play the clarinet. I am interested in the musical practices of unknown and underrepresented cultures and committed to promoting them in research. That is why I am studying the talking drum of the Yorùbá. In 2020, my Ph.D. project on the dùndún talking drum and how it makes people feel won the African Studies Review competition for the best African-based dissertation. I love making music and doing sports during my free time.



TINA ROESKE

I have just become a teacher after 15 years of research in Germany and the US. My main interest was birdsong—the “music” of birds! I wanted to find out how what birds do when they sing is similar to what humans do when they make music. I liked the many parallels that I found—they suggest that the reason we produce beautiful music might have little to do with us being a smarter, more advanced species. Perhaps getting pleasure from sounds is all you need to develop “music.” Now I enjoy discussing questions like these with my high school students.



MELANIE WALD-FUHRMANN

I am the director of the music department in the Max Planck Institute for Empirical Aesthetics. As a kid, I was already deeply interested in music, its forms, meaning, and effects on people. So, I became a musicologist. After my studies and a Ph.D. that I obtained in Zurich, I was a professor of musicology in the German towns of Lübeck and Berlin before starting at my current position.



PAULINE LARROUY-MAESTRI

I am a researcher at the Max Planck Institute for Empirical Aesthetics. I knew from the start that my work would be related to music but instead of becoming a concert pianist, I use music to better understand how we process (and enjoy) sounds. After studying music therapy, pedagogy, psychology, music, and speech-language pathology, I obtained my Ph.D. (in 2013) and visited several laboratories across the globe. Now, I develop research at the intersection of these different interests and currently focus on the formation of perceptual categories in both music and language.