

# These auditory illusions make people hear phrases as melodies

A linguist explains why spoken words sound melodic when repeated over and over.

BY BEN MACAULAY | PUBLISHED JUN 22, 2023 9:00 AM EDT

---

*In [Head Trip](#), PopSci explores the relationship between our brains, our senses, and the strange things that happen in between.*

**CAN YOU TELL** when you've suddenly and accidentally broken into song? When you say, "I don't know," does it sound like a melody that you can hum? Why does a spoken word, when repeated over and over, begin to sound musical?

To understand why people sometimes hear uttered phrases as melodies, let's first break down language to its auditory elements. Not vowels and consonants—that's too far. Instead, let's focus on what linguists like me call prosody, or the intonation, stress, and rhythm of individual syllables.

In 1995, Diana Deutsch, an expert on [musical illusions and paradoxes](#), coined the term *speech-to-song illusion*. She was reviewing a longer recording of herself and found that the seemingly random phrase *sometimes behave so strangely*, clipped out of context and left on repeat, began to sound like a song, complete with rhythm and melody. (You can hear [a loop of the recording](#) online.)

As it turns out, the illusion does not stem from Deutsch's voice being especially musical. This illusion happens with all sorts of spoken stimuli that are played at regular intervals. The effect is most intense when the repeated stimulus is in a language that is very different phonologically from the listener's native language. For example, in one 2015 study, [English speakers most easily heard spoken Irish as song](#), followed by Hindi and Croatian.

Why is repetition so important here? There are measurable differences in how verbal stimuli sound when played once versus on repeat. In a 2008 study by Deutsch and colleagues, [31 participants were asked to repeat the stimulus themselves](#). After hearing it once, they spoke the sequence of words they'd heard; after 10 repetitions, they began to sing their responses, with a slightly larger pitch range than the original stimulus and the spoken responses. What's more, according to a 2013 fMRI study (by a team that included Deutsch), [repeated stimulus activates a part of the brain](#) that is also activated by song and associated with the processing of complex pitch patterns, a network of eight different regions.

While Deutsch is credited for discovering the illusion in the realm of psychology, the basic premise—that repetition will add a layer of musicality to a recording of speech—was recognized long before 1995. Specifically, her use of a clipped recording on repeat reminded me of sampling, a technique used in hip-hop in which clips of existing recordings (sometimes speech) are incorporated in a new musical arrangement.

To put this practice into context, I spoke to two experts: Langston Wilkins, expert in hip-hop and assistant professor at the University of Wisconsin at Madison, and Dan Charnas, historian of hip-hop and associate arts professor at New York University. Both confirmed that the use of repetition to add musicality to spoken vocal samples is a common practice in hip-hop, but neither was familiar with Deutsch's framing of the phenomenon as an auditory illusion. Wilkins describes some of the sources hip-hop takes spoken samples from, including sermons, speeches, and film and TV dialogue. Charnas notes that the practice actually precedes hip-hop and tape recorders, with many vocal traditions originating in the African American community and gaining their musicality through repetition, for instance rap. The fact that these traditions were pioneered by African Americans may explain how the speech-to-song illusion

evaded academic studies of musicality until the 1990's—as Charnas notes, this reflects a general trend in academia failing to center African American culture, a culture in which “speech and song...have always been twin.”

But why does the illusion happen? As a linguist who studies intonation, I see a number of factors that may be at play here. First, I think it is significant that Deutsch and other scholars use clips cut out of longer recordings to demonstrate the illusion. Sentences in English and other languages have intonation, which includes pitch movements and other acoustic patterns that can span a full sentence. Cutting words out of a sentence will make a lot of these patterns incomplete. In a way, it is like cropping *aterm* out of the word *watermelon* and finding that it stops sounding like a word. The intonation that remains on Deutsch's *sometimes behave so strangely* doesn't sound like that of a full sentence. But unlike sentences, songs can have any melody they want. Perhaps that makes it easier for us to process phrases and clips like these as singing.

Another way to break down the speech-to-song illusion is by thinking about how humans process the “melody,” or pitch contour, of speech separately from the consonants and vowels that it is realized on. One example from English is the phrase *I don't know*, which can be condensed to *iunno*, or even hummed as the melody you would expect to hear on *iunno*: a rise-fall-rise pitch contour. The hummed version of this melody contains none of the consonants or vowels in the phrase *I don't know*, but people still know it means *I don't know*, especially when you pair it with a visual cue like a shrug.

The process is not unique to English, either. One of the languages I studied for my dissertation, Amis (an Indigenous language of Taiwan), has its own *iunno*. In Amis, the phrase *i saw* (pronounced like “ee sow”) is added for emphasis on phrases like *Really? You can either say the phrase in full, or you can just invoke the tonal melody that would have been on the phrase.* These examples show that humans who are not hard of hearing can home in on the melody of speech, a key component of the speech-to-sound illusion.

“The illusion is weaker in speakers of tonal languages,” says Andrew King, director of the Centre for Integrative Neuroscience at the University of Oxford. He's referring to languages like Chinese and Yoruba in which tones are an integral part of the phonetic structure of words, alongside consonants and vowels. This tells me that the ability to separate the tonal contour from the consonants and vowels must be a crucial component of this illusion.

When asked about people who might pick up on different aspects of the speech signal due to differences in hearing, Neil Bauman, founder of the Center for Hearing Loss Help, immediately saw a connection between the speech-to-song illusion and audio pareidolia, a phenomenon in which speech or music are perceived in random sounds not produced by humans. Bauman explained that like the speech-to-song illusion stimuli, some cases of audio pareidolia do involve a rhythmic element, such as a fan with a loose bearing that creaks and whizzes at fixed intervals, which may more easily trigger an audio illusion.

While scholars are still working on exactly how and why the speech-to-song illusion happens, the fact remains that under the right conditions, people can hear songs when none were intended. This finding has been replicated, on repeat, both in the laboratory and in the DJ booth.

*Ben Macaulay is a lecturer in English linguistics at Lund University. His research focuses on prosody and intonation, namely the production, processing, development, and documentation of sentence-level tonal contours in the world's languages. He received a Ph.D. in 2021 at the Graduate Center, CUNY, and his dissertation project was a typological study of intonation in the endangered Indigenous languages of Taiwan based on novel fieldwork.*

Read more [PopSci+ stories](#).

[PSYCHOLOGY](#)

[HEADTRIP](#)

[MEMBERSHIP](#)