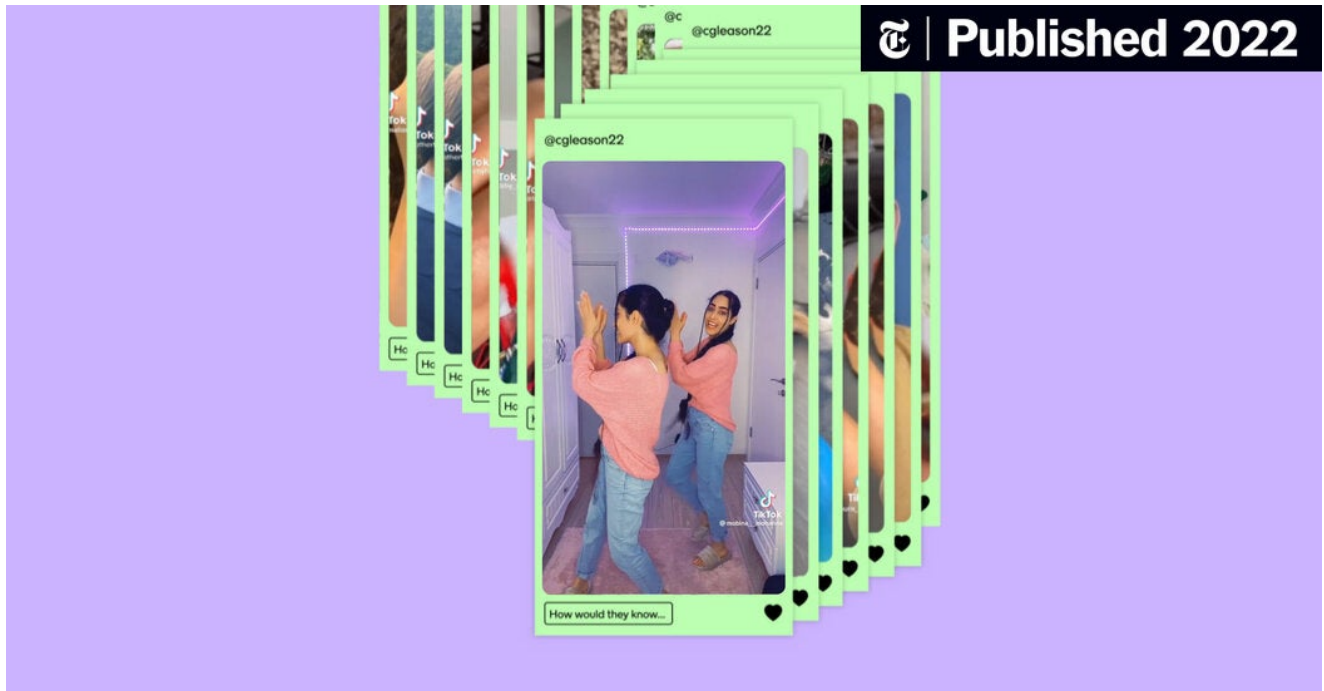


TikTok Audio Memes Are Everywhere. How Do They Work?

 [nytimes.com/interactive/2022/08/17/magazine/tiktok-sounds-memes.html](https://www.nytimes.com/interactive/2022/08/17/magazine/tiktok-sounds-memes.html)

Charlotte Shane

August 17, 2022



Why Do We Love TikTok Audio Memes? Call It ‘Brainfeel.’

By Charlotte ShaneAug. 17, 2022

On March 25, 2020, Chris Gleason was in bed at his parents' house in Pennsylvania, thinking up ideas for videos that might go viral. Just before graduating from college with a musical-theater degree in 2019, he took a job at a nautical-themed restaurant in the Washington, D.C., area, where he served oysters and cocktails with names like Boston Tea Party and Blown Off Course. When Covid-19 temporarily shuttered indoor dining, he quit and moved back home before attending business school. In the interim, he recorded two or three videos a day, writing scripts and editing the footage on his phone. Then he uploaded the results on TikTok.

That month, in the early days of the pandemic, American adults spent well over a billion hours on the platform, which had become the most downloaded nongame app in the world. A few of Gleason's posts — him dancing to the "Law & Order" theme, a skit about clueless restaurant patrons — had gone modestly viral in the past, and he was intrigued by the possibility of making a megahit. TikTok had given so many users their 15 minutes of fame. Surely he, with his performance background, could be among them. What he came up with — a mocking take on his conflicted inner dialogue — is now cultural history.



The post has been viewed more than 14 million times, but the reach of its exasperated exchange — *Nobody's gonna know. They're gonna know* — is much, much larger. When a creator uploads a video to TikTok, they have an option to make that video's audio a "sound" that other users can easily use in their own videos: lip-syncing to it, adding more noise on top or treating it as a soundtrack. Gleason's sound has been used in at least 336,000 other videos.

Through that repurposing, Gleason, who now works in advertising in New York, has gone viral again and again. Footage of a lone tourist climbing to the top of Chichén Itzá in Mexico has been viewed 72 million times; a restaurant's demonstration of how you can cut a whole pizza to disguise eating a slice, 82 million. This year, the actress and model Shay Mitchell used the sound when she announced her second pregnancy, following in the steps of the singer Meghan Trainor, who used it in 2020 when she was in the third trimester of her first pregnancy.

Gleason's dry delivery, coupled with the instrumental score he discovered while searching for dramatic reality-TV-show tracks, turned out to be ideal meme material. Generic enough to apply to whatever scenario in which viewers might find themselves, it combined high-stakes drama and spot-on comic timing. Plus, it's short. "I tend to be a little long-winded," Gleason said while reflecting on his near-instant classic. "But that one worked out to be 22 seconds." (The accompanying score, named "Primal Fear," was released by Dave James in 2011, and thanks to Gleason's boost, leads a robustly meme-ed life of its own.)

Gleason's voice, more than Gleason himself, is the star: The original post's comment section is still frequented by people expressing shock that they've finally found the source after tracing it through its reuses. (Often, they say they were convinced that the dialogue was from an actual reality-TV show.) Millions of people know how Chris Gleason sounds but have no idea what he looks like. "Whenever I'm out with my friends, they're like, 'Oh, Chris is famous,'" Gleason said. "But I don't feel famous. Because people only know my voice."

Welcome to the era of the audio meme, a time when replicable units of sound are a cultural currency as strong as — if not stronger than — images and text. Though TikTok didn't invent the audio meme, its effortless interface may have perfected it, and the platform, which recently ended Google's 15-year-long run as the most visited website in the world, would be nothing without sound.

And what a range of sound there is. TikTok is well known as a music-industry hitmaker integral to the success of pop stars like Doja Cat and Megan Thee Stallion, as well as sleekly produced artists with writing teams capable of engineering the catchy, often danceable hooks that blaze through the app. Homemade covers — someone singing in their bedroom a cappella or accompanied only by keyboard or acoustic guitar — can get traction, too. But "the viral canon" is made up of much stranger sounds: evocative line readings from TV and film, a child beatboxing, an amateur golfer swearing. The teenage user @couchtable's "accent challenge" — a mushy, nearly unintelligible recital of slang with a hyperexaggerated Southwest Missouri accent — has reached tens of millions of viewers. A clip of a video-game character's echoey shouts of "Hoo! Hah! Oi!" was renamed "WHY IS EVERYONE USING THIS" after it served as the soundtrack for extremely popular puppy videos, boyfriend-girlfriend skits and a sendup of a bikini barista's pervy customers.

Why are we drawn to such uncategorizable sounds, the noises that deliver limited-to-no-information yet elicit our adoration? If “mouthfeel” is used to indicate the visceral experience of consuming food and drink, “brainfeel” might be a decent descriptor for what makes a sound compelling beyond musical qualities or linguistic meaning — though the sensation hits within music and language, too. A funny pronunciation that you can’t stop imitating, the drop that gets the whole club jumping, the plaintive meow of a cat, the key that turns in your heart when you hear someone speak with great emotion: That’s brainfeel, ineffable and affecting and addictive.



Older meme-generating hotbeds like Twitter, Reddit and 4chan rely on silent, visual communication. And while it isn’t exactly labor intensive to type text over a still from “The Simpsons” or plug it into the empty panels next to Drake dancing in the “Hotline Bling” video, you still have to pull the image, open a program to tamper with it, then move it to wherever you want it. Using an uploaded sound on TikTok takes a few taps, and you never leave the app.

This functionality traces back to TikTok’s 2018 merger with Musical.ly, another Chinese-owned video app, one focused on lip-syncing. According to internet lore, what became TikTok’s sound feature was known on Musical.ly as “remuse” (instead of “reuse”). One way or another, the function created an unprecedented mode of cross-user riffing and engagement, like quote-tweeting for audio. Occasionally, TikTok delivers a piece of viral content in which the visuals can’t be parsed from the sound. Nathan Apodaca, @420doggface208, may have created the blueprint for this when he recorded himself skateboarding on a sunny day in September 2020, drinking from an Ocean Spray bottle and lip-syncing along to Fleetwood Mac’s “Dreams.” But much more often, TikTok virality, and its ability to create culture that travels off the app, depends on memeifying sound.

Before social media, Gleason’s “Nobody’s gonna know” might have been called a catchphrase: a banal word combination animated by unique context and delivery. “Did I Do That?” “I’ll Be Back” and “How You Doin’?” would mean nothing if not for the precise tones and cadences with which their originators (Jaleel White as Steve Urkel, Arnold Schwarzenegger as the Terminator and Wendy Williams as herself) so reliably rendered them. In a phone call, the linguist Molly Babel mentioned Alicia Silverstone’s “As *if*,” from the movie “Clueless”: Taken altogether, Silverstone’s iconic phrasing, intonation and cadence *are* the sound. Like earworms, these quips are so mentally sticky that it takes just a few listens for your mind to latch onto them and never let go. Try reading them without hearing their corresponding acoustic signatures in your head: “Here’s Johnny!” “You talkin’ to me?” “Damn, Daniel!”



“Memes are often symbols,” says Don Caldwell, editor in chief of the dizzyingly comprehensive website Know Your Meme, and exceptionally viral memes tend to be “very novel or very catchy or just very, very striking emotionally.” Even when they’re estranged from their origins — i.e. taken out of context — they’re funny or moving or both. He mentions “sad trombone” as a pre-internet audio meme, and it occurs to me that the song “Yakety Sax” counts, too. Both musical cues evoke an unmistakable mood in and of themselves, but after decades of application to that effect, their deployment adds another layer of information to whatever scene they orchestrate. It’s a wink to the audience that positions the moment within a cultural continuum. The famous Wilhelm scream, a histrionic stock effect taken from a 1951 film, has since appeared in more than 100 movies, where it has become an inside joke for sound engineers and film fans. An audio meme’s most crucial quality, though, is the ability to instantly excite us, to make us think, upon the first listen: *I need to hear that again.*



The Brooklyn native Joel Joseph, known online as Lord Hec, had amassed about 200,000 followers by September 2021 when “Love Nwantiti,” a mellow, haunting song by the Nigerian singer CKay, exploded on TikTok. Several influencers choreographed challenges for the song, but Joseph, a 24-year-old dancer and instructor who has been creating content online for almost a decade, got hooked on a smooth, playful version set to the pre-chorus. One morning, while in Las Vegas for work, he recorded a vocal track in his hotel bathroom to go with a performance that he shot later that day by the side of a large backyard pool. The audio consists entirely of exuberant cues and hype noises (“Jump! Then you gotta bend — point! Hey, hey, hey, clap clap!”) used to keep time with the music. And his delivery is so confident and joyful that the visuals of the dance almost become secondary to the sonic experience of his personality.



“When I teach my students, I make these sounds instead of doing 5-6-7-8, the typical count,” he said. “It’s easier for me to remember each thing by either stating what’s happening or making a sound associated with it.”

Joseph hadn’t set out to make a viral sound. It was, after all, simply intended to teach viewers the dance, and he didn’t expect people to consider it separate from the visual. But once [the video](#) was uploaded, his fellow Tiktokers bombarded him with a request that often appears on the app: “Make this a sound.” The excerpt from “Love Nwantiti” [used by @itsjustnifee](#), the dance’s originator, currently has 689,000 uses. Joseph’s “dance lessons” version, which includes his punctuating sounds, has 1.5 million.

Joseph’s tutorial sounds aren’t quite music. They vary in pitch to create emphasis, and he keeps a rhythm, but he isn’t singing. Nor, however, is he speaking in the conventional sense, because not everything coming out of his mouth is a word. (“Love Nwantiti” itself has quite a

few nonword lyrics; the chorus consists entirely of “ah” repeated several dozen times.) These silly noises were what pleased people the most, judging by the comment section; that can happen offline in the classroom, too. “My students giggle,” Joseph says, “but then they start saying the same thing. And if I only played the song, they would say, ‘Can you make the sounds?’” Just as Joseph tried to describe his motions with his vocalizing, so did delighted TikTok users try to approximate those sounds with their stylized, phonetic spellings: “SoLo lo~” “A- A- A-” “then your gonna network a a i i i i i”

This is how users share the pleasure of an audio meme in the silent space of typed comments. Creative phonetic renderings — attempts to convey the brainfeel of a sound — are all over TikTok, especially when the sound in question involves a human voice. When I sent some examples to Babel, the linguist, to find out how accurate they were, she was impressed. In a video by Caitlin Reilly that mocks insipid wedding vows delivered with maximum vocal fry, Babel noted that commenters tried to capture the timing of the speech by giving syllables prominence. “Jason” is “jeey-senn,” and “today” becomes “TaHdayH.” Another sound, which accompanies footage of a cat writhing on sunny pavement, consists of the creator @owlfacexd cheerfully testing out pronunciations of the word “concrete,” stylized as “conkcrente !” in the video’s caption. (Naturally, commenters ran with the theme, offering up “comcrente” “comkrete” and “CONK CRETE!!!!” among others.) For these, Babel praised what she called the orthographic rendering, surmising that the m’s indicated a nasal sound.

Babel studies vocal attractiveness, in part because the existing studies she came across early in her academic career were methodologically limited. Psychologists were trying to divorce voice from language by having a speaker do something “robotic,” like sustain a single vowel sound, which removes the special aspects of language and could result in someone self-consciously adjusting their voice in an unnatural way. (These papers also tended to situate attraction exclusively in “sexual space,” as if we don’t enjoy the voices of children or grandparents or whiny, lovable nerds like Steve Urkel.) Babel and her collaborator, Grant McGuire, found an affinity for voices that “recapitulate gender stereotypes” — meaning men who sound larger and women who sound smaller, for instance — in part because we like predictability and familiarity. But “there’s an attentional draw to voices that are atypical to us.”

I asked if it would be accurate, then, to say that we like unusual voices or that we like unusual voices only if the content of the speech is intelligible. “These are top-notch research questions that we still don’t really have answers to,” she said. But she felt confident saying that “we just like variability sometimes. We want to hear a little bit of novelty, we might want to hear a little bit of modulation in pronunciation” because it helps keep our attention. We also like information, and expressive voices give us still more to process — even when (or perhaps especially when) they’re making sounds instead of words.

The tension between predictability and novelty comes up a lot with sound. Predictive coding — a theory that holds that our brains make predictions about what the next element in an unfolding pattern will be — is a crucial element of music, the neuroscientist Robert J. Zatorre told me. In fact it's a big part "of all cognition" that our brains are "constantly figuring out what might happen next." Our reward systems engage when we listen to music, based on previous experience. If we hear what we've learned to expect without any deviation, no dopamine is released. If we hear an alteration that was hard to predict, we might get a dopamine boost. But if we don't hear something that we knew to expect — because a musician hits an unintended note or your navigation app interrupts a song's climax — our dopamine level drops. "Your system actually gets inhibited," Zatorre explained.

Something interesting happens, though, when the expectation is not only met but exceeded: That gives us a huge dopamine burst. This could explain what happened with Joseph's dance lesson. TikTokers knew the base song very well and could still hear the track accompanying the dance. But they got Joseph's happy vocalizing on top: something they weren't expecting combined with something they were. After enough repeat listens, Joseph's vocal track became its own separate phenomenon, almost its own song, something fans could sing without "Love Nwantiti" playing underneath.

Audio-meme magic is unpredictable and, at the same time, feels obvious and inevitable after the fact. Once you've heard the sound, while you're hearing the sound — the Missouri-patois parody, the breathy hoots of a video-game hero — you hear that it's wonderful, irresistible.

That was true for the impromptu serenade of a neighborhood cat named Mashed Potatoes that @june_banoon, a teacher currently living in South Korea, posted in the summer of 2021. "That was the most aimless singing I've ever done in my life," @june_banoon told me. "I used to sing opera in high school. I used to sing in competitions. So for that little bit of complete aimless, pointless singing" to go viral "was astonishing to me." People around the world really like cat content. And people around the world, overcome with appreciation for a little animal they like looking at, sing to their cats all the time. But June's voice, characterized in comments as "angelic" and "like a Disney princess," paired with the elegant simplicity and accuracy of the lyrics, achieved the platonic ideal of a pet tune. "Here comes the boy," June sings as Mashed Potatoes leisurely waddles toward the camera. "Hello, boy. Welcome. There he is. He is here." June told me: "It was one of those things I originally anticipated posting and deleting within an hour in case nobody really liked it." It currently has more than 42 million plays.



"We humans have two major auditory communication systems," Zatorre said. "One of them is speech, of course — language. But the other is music." And music, in fact, precedes speech: "Parents sing to their infants in every single culture. Lullabies exist in every culture." Music,

like food, activates the circuitry of our neurological reward system, which exists to compel us toward the most necessary elements of survival and so shapes our behavior from the earliest age. We don't need music to survive, Zatorre says, and yet we're clearly driven to seek it out because of how it affects us.

TikTok users have confessed intimate details to June — that the clip reminds them of a recently deceased parent or that it helps them sleep well. I've probably heard "Here Comes the Boy" a hundred times, and yet as I recalled the clip while typing out its lyrics, I teared up. If you asked me why, the only explanation I can offer would be brainfeel.

The power of music, Zatorre says, comes from the neurological pleasure it gives us and, more broadly, "from the emotional engagement we get." Music generates social bonds and so is related to empathy, the ability to connect to another person. Connections occur on TikTok when creators duet each other's videos (posting their new recording side-by-side with a pre-existing one) to add another layer of sound or savor and trade phonetic spellings in the comments, and those attachments can be lasting. What happens on the app doesn't stay on the app, which is why it's such a formidable cultural force — and a strong interpersonal one.

"I was recognized in the subway the other day by someone who recently binged my whole account," June said in a speaking voice as mellifluous as singing. "We went out to lunch."

Charlotte Shane is a writer and an editor who co-founded the Brooklyn-based TigerBee Press in 2015. She is the author of the epistolary memoir "Prostitute Laundry."

Additional design and development by Jacky Myint.