“A Lesson In The Language Of ‘Baby Talk’” Excerpt Transcript

Excerpt from October 13, 2017 episode of Science Friday.

JOHN DANKOSKY: OK, imagine you're sitting on a train, and you overhear a woman sitting behind you saying this.

PERSON 1: Hi! Hi! What are you looking at?

DANKOSKY: You could probably tell, even without seeing the woman, that she is speaking to a baby-- we hope. Now, baby talk isn't just some adorable or irritating affectation that we adults acquire when we're around infants. It actually gives babies important acoustical information that they need to begin to process words in our language.

But new research out this week suggests there's something else adults are doing when they use baby talk. They change the timbre of their voice as well. Here with me to discuss this is Elise Piazza, post-doctoral researcher at the Princeton Neuroscience Institute at Princeton University. Welcome to Science Friday.

ELISE PIAZZA: Hi. Thanks very much for having me.

DANKOSKY: Why don't you explain what timbre is first, so people understand what we're talking about.

PIAZZA: So timbre is generally defined as the unique quality or the tone color of a sound. And it's actually much less understood than pitch or rhythm, but we rely on it constantly to distinguish and enjoy all of the different flavors of sounds around us.

So for example, we can easily discern different idiosyncratic celebrity voices, like Barry White, who has a famously velvety voice, or Gilbert Gottfried who has a much more nasal voice, or maybe Tom Waits with his sort of gravelly growl, even if these three people were all singing the same note with the same rhythm. And if you imagine an orchestra tuning up, they often will all play the same pitch, like A 440, but you can still easily pick out the different instrument families or textures throughout the orchestra, including the reedy woodwinds, the buzzy brass, the mellow strings, et cetera. And these are all timbre descriptors.

And timbre can also give us the overall gist of sounds, in a way that many other features really can't. So research has shown that people can identify the genre and the era of a piece of music quite well, in less than half a second. And this is probably because timbre is giving us some clues about the instrumentation and the texture of that music.

DANKOSKY: OK, we found that mothers do this regularly during baby talk, of course-- dads, too. We've got a clip here of a mother speaking to her baby, and then speaking in her normal voice. I'm going to try and guess the timbre differences. First, let's take a listen.

PERSON 1: Hi. What do you see? You see a microphone? Yeah?

Anything shiny catches her attention right now.

DANKOSKY: OK, so that's a mom and a baby. So I heard sort of a breathy quality to this, a breathiness. But what else do you hear in there, Elise?

PIAZZA: So some of the features of motherese, or infant-directed speech, that have already been well-documented in the literature which you can easily hear in that segment, include the much more exaggerated pitch contours, so the sort of swoops up and down. Also, the overall pitch is much higher, which we think babies tend to prefer. And the repetition and the rhythm is also quite salient, so there are more pauses in the motherese. All of these cues sort of combine to help to highlight the inherent structure in speech, so to help babies to segment this constant stream of noise into the building blocks of language, like syllables, words, and sentences.

DANKOSKY: You talked about timbre helping us to get at the gist of sounds. What exactly do you mean by that?

PIAZZA: So if you take a really quick slice of music, you're not getting much information about the rhythmic complexity. So you might not be able to hear the beat structure yet, and you might not know which notes, or sort of which key center-- whether you're, let's say, in a major or a minor mode-- just from a second of music.

But what you can hear are things like the instrumentation, whether we're dealing with the timbres of a rock band, maybe the sort of percussive sounds and the sort of guitar timbre, as opposed to maybe a more chamber music kind of timbre. And similarly, with the voice, you can really get a lot of information about someone's identity from pretty short clips of voices, which you might not be able to just by a single note or a single piece of rhythm.

DANKOSKY: Well, one of the most interesting findings here is about whether or not this happens across all languages, these qualities of timbre. What did you learn about that? Is it just in English or not?

PIAZZA: So we found statistical commonalities in the exact structure of the shift that generalized across a very wide variety of languages.

We first brought the English-speaking mothers into the lab, and we asked them to just speak as they naturally would to their babies while playing with them, while reading them some age-appropriate board books. And then we had them speak to an adult experimenter. And we used a sophisticated machine-learning algorithm, just sort of a form of AI, where we basically designed a model of this shift in timbre between infant- and adult-directed speech in the English-speaking mothers.

And the most remarkable thing was that when we then brought this new group of 12 mothers, also from the central New Jersey area, who spoke this wider range of languages from around the world, the same model that we had devised to discriminate these two modes in the English group generalized immediately to this new group of non-English speakers. And we used these rigorous statistical methods to determine that. So for instance, if you train the model just on the English data, it actually generalizes extremely well to the non-English and vice versa.

DANKOSKY: It's fascinating. I didn't know we could learn this much about baby talk. But I'm very appreciative that you took some time with us today. Elise Piazza, post-doctoral researcher at the Princeton Neuroscience Institute at Princeton University. Thanks so much.