

[SQUEAKING]

[RUSTLING]

[CLICKING]

NANCY

KANWISHER:

All right, OK, so let's start. We're talking about music today, which is fun and awesome. But first, let me give you a brief whirlwind reminder of what we did last time. We talked about hearing in general and speech in particular. And we started, as usual, with computational theory, thinking about what is the problem of audition and what is sound. It's the first step of that.

And sound is pressure waves traveling through the air. And the cool thing about hearing is that we extract lots of information from this very, very simple signal of pressure waves arriving at the ear. We use it to recognize sounds, to localize sounds, to figure out what things are made of, and to understand events around us, and all kinds of things.

And these problems are a major computational challenge. And in particular, they are ill-posed. That means that the available information doesn't give you a unique solution if you consider the computational problem narrowly.

And that's true for separating sound sources. So if you have two sound sources at once, say, two people speaking or a person speaking and a lot of background noise, that's known as the cocktail party problem. Those sounds add on top of each other. And there's no way to pull them apart without bringing in other information, knowledge about the world or knowledge about the nature of voices or speaking or who's speaking. Or you need something else, or else it's ill-posed. That is not solvable just from the basic input.

Another case of an ill-posed problem in audition is the case of reverb. So the sound that I'm making right now that's coming out my mouth is bouncing off the walls and is arriving at your ears over each little piece of sound that I make is arriving at different latencies after I say it as it travels different paths bouncing around the room. There's not too much reverb in here, so it's not that noticeable. But if we did this in a cathedral, you'd hear all these echoes.

OK, and so that makes another ill-posed problem, because all of those different sounds are added on top of themselves diminished in volume over time. And you get the sum of all of those, and you have to pull it apart and figure out what that sound is.

So both problems are solved by using knowledge of the real world. In the case of reverb, it's actual implicit knowledge that you all have that you didn't know you have about the physics of reverb. Because if we play you sounds with the wrong physics of reverb, you won't be able to deal with reverb. And that says it's implicit knowledge in your head, which is pretty cool, that you use to constrain the ill-posed problem.

We talked about speech. Phonemes are sounds that distinguish two different words in a language, like make and bake. Those are two different sounds that make the difference between two words. Each possible speech sound is not a phoneme in every language of the world. Languages have some subset of the space of possible phonemes that distinguish words in their language. Phonemes include vowels that have these stacked harmonics in the spectrogram, and consonants which are the quick transitions in the vertical stripes in the spectrogram, leading into the harmonic stacks of vowels.

We talked about the problem of talker variability, that a given phoneme or word sounds very different, looks very different in the spectrogram if spoken by two different people. And conversely, the same person speaking two different words looks very different in the spectrogram. And so that means that the identity of the speaker and the identity of the word being said are all mushed up together.

And that means that if you want to recognize the voice independent of what's being said, or recognize the word independent of who's saying it, you have a big computational challenge, a classic invariance problem. Yeah, Ben.

AUDIENCE: I don't mean to hold us up. I just wanted to make sure that I'm understanding. So the difference between consonants and vowels, are vowels just harmonic, like connective elements between consonants? And are consonants the percussive? Or are they actual-- like, I just didn't understand that.

NANCY
KANWISHER: Yeah, so in the spectrogram, those-- I didn't put that on the slide here-- but those horizontal red stripes in the slides that I showed you last time, those in the spectrogram, those are bands of energy at different frequencies that are sustained over a chunk of time. And those are typical of vowels, or singing or musical. And those harmonic sounds that have pitch. And so vowels have those sustained chunks that look like this in the spectrogram. And then there are these weird vertical stripes and transitions in and out of the vowels that are the consonants.

AUDIENCE: Vowels are when you don't have [INAUDIBLE] spectrographs because air is just flowing through and you're filtering it somehow, like positioning your vocal tract in a certain way. And consonants are when you close off that air or restrict it in some way. So like S's and F's, you're not closing all the way off, but you're really constricting the vocal tract. And in a lot of other consonants, you're actually fully closing it.

NANCY
KANWISHER: OK, and then we talked a bit about the brain basis. And I pointed out that the neural anatomy of sound processing-- the subcortical neuroanatomy is much more complicated than the subcortical neuroanatomy of vision. In vision, you have one stop in the LGN, and then you go up to the cortex coming up from the retina.

In audition, you have many stops between the cochlea, where you pick up sounds in the inner ear, and auditory cortex. Some of those stops are shown up here. And we didn't discuss them.

So then we talked about primary auditory cortex. That's on the top of the temporal lobes, like right in there medially. You went in. And it has this tonotopic property, and that is a map of frequency space with this systematic high-low-high mapping of frequency space that you can see here-- high, low, high, like that. This is the top of the temporal lobe right there.

And I pointed out that in animals and in one recent MRI study, the response properties of primary auditory cortex are well modeled by these fairly simple linear filters, known as spectrotemporal receptive fields or STRFs, shown here. So they're simple acoustic properties of a given band of frequencies rising or falling at different rates.

So today, we're going to talk about music. And this is also an important moment in the course. Because up to now, we've been talking about functions that are mostly shared with animals. Speech is kind of on the cusp.

I was going to make this point before speech. And that's actually muddy, because lots of animals are really good at speech perception. Chinchillas can distinguish ba from pa. Go figure, anyway. So they can perceive speech, but obviously they don't use it in the same way.

But music is most definitely uniquely human. And so most of the things we'll be talking about from here on out are things about the human brain, in particular. And I think these are the coolest things in human cognitive neuroscience, because they tell us something about who we are as human beings.

But they are also the hardest ones to study. Why is that?

AUDIENCE: [INAUDIBLE]

NANCY
KANWISHER: No animal models. And I'm always lamenting how-- about the shortcomings of each of the methods in human cognitive neuroscience. And we have lots of them, and they complement each other, but there's a whole host of things that none of those methods are good for.

And so now we're really out on thin ice trying to understand these things with a weaker set of methods where we can't go back and validate them with animal models. And that's just life. That's what we do.

So now let's back up for a second and consider, why am I allocating a whole lecture for such a fluffy, frivolous topic as music. And I would say, that's because it's not fluffy. It's actually fundamental.

And it's fundamental in the sense that music is both uniquely human-- no other animal has anything remotely like human music-- and it's also universally human. That is, every human culture that's been studied has some kind of music.

So music is really an essential part of what it means to be a human being. It's really at the core of humanity. And that alone makes it interesting. But further-- question?

AUDIENCE: So, like, birdsong--

NANCY
KANWISHER: Birdsong doesn't count. No, birdsong doesn't count in all kinds of ways. One, it doesn't have anywhere near the flexibility and variability. There are like narrow domains in which each male zebra finch makes a slightly different version of the call, but within an extremely narrow range.

There's actually a brain imaging study that looks in brain imaging in songbirds and asks, do they have reward brain region responses to music. And the answer is, yes, in some cases. Like, do they enjoy it, right, is that part of-- and the answer is yes, but only when the significance of the birdsong is something that's relevant to them, like, there's a potential mate right here, then they like it.

But they don't like it just for the sound. And that makes it very different from humans. And there are other differences as well.

So it's further really important to us humans in a whole bunch of ways. One, we have been doing it for a very long time. And so, for example, the archaeological record shows these 40,000-year-old bone flutes that you can see from the structure of flute make particular sets of possible pitches. And further, most people who've thought about this have argued that singing probably goes back much farther than the bone flutes.

After all, you don't have to make anything to do it. You can just sing. Some have even speculated that singing evolved before language. It's just speculation, but that's possible.

In any case, it goes way back evolutionarily. It also arises early in development. So very young infants are extremely interested in music. They're sensitive to beat and melody, independent of pitch. We'll talk more about that a little bit.

And finally, if you're not impressed with any of those arguments, people spend a lot of money on music. And if that's your index of importance, it's really important. Last year, \$43 billion in sales.

So I'd say it's not a frivolous topic. It's a fundamental topic. It's near the core of what it means to be a human being. And all of this raises a really obvious question. Why do we create and like music in the first place? What is it for?

And this is a puzzle that people have thought about for at least centuries, probably millennia. And this includes all kinds of major thinkers, like Darwin, who said, "As neither the enjoyment nor the capacity of producing musical notes are faculties of the least direct use to man in reference to his ordinary habits of life, they must be ranked amongst the most mysterious with which he is endowed."

So Darwin is implicitly assuming here that music is an evolved capacity. It's not something that we just learn and that cultures invent, if they feel like it or don't feel like it. But it's actually evolved and shaped by natural selection. And that means there must be some function that natural selection was acting on that was relevant to survival.

So people have speculated about what that function might be. Those who think that music is an evolved function, including Darwin, he speculated that it's for sexual selection. And his writing is so beautiful, I won't paraphrase it.

He says, "It appears probable that the progenitors of man, either the males or females or both sexes, before acquiring the power of expressing their mutual love in articulate language, endeavored to charm each other with musical notes and rhythm." So that's Darwin's speculation. It's just a speculation, but a lovely one.

Also, note that he threw away this radical idea in here: "before acquiring the power to express their mutual love in articulate language." So he's speculating that music came before language. Again, all speculation, but interesting speculation.

More recently, up the street, there's a bunch of people who've been thinking about this a lot. And Sam Mehr at Harvard has been arguing that the function of music and song, in particular, which he thinks is really the fundamental basic kind of native form of music, has an evolutionary role in managing parent-offspring conflict. And that's something that many evolutionary theorists have written about.

The genetic interests of a parent and an offspring are highly overlapping, but not completely overlapping. The parent has other offspring to take care of besides this one right here. That one right there wants 100% of the parent's effort. Therein lies the conflict.

And so Mehr has proposed that infant directed song arose in this kind of arms race between the somewhat competing interests of the parent and the offspring. And it manages this need the infant has to know the parent is there with the fact that the parent has other needs, so i guess idea they can sing while attending to other offspring, and on and on. So there's other kinds of speculations like this.

But importantly, this is not the only kind of view. It's not necessarily the case that music is an evolved capacity. So others have argued that it's not.

So Steve Pinker, also up the street, has argued that music is "auditory cheesecake, an exquisite confection crafted to tickle the sensitive spots of at least six of our mental faculties. If it vanished from our species, the rest of our lifestyle would be virtually unchanged." I think that might say a little more about Steve Pinker than it does about music. Nonetheless, it's a possible view.

What he's saying is that music is not an evolutionary adaptation at all, but an alternate use of neural machinery that evolved for some other function. And then once you have this neural machinery, what the hell, you can invent cultural forms and use it to do other things like music.

And the most obvious kind of neural machinery that you might co-opt for that function would be neural machinery for speech or neural machinery for language, which, as I argued briefly last time, are not the same thing. One is the auditory perception of speech sounds and the other is the understanding of linguistic meaning.

So the nice thing about this is, finally after all this entertaining but speculative stuff, we have an empirical question. This is something we can ask empirically. Does music actually use the same machinery as speech or language, or does it not? Some of the rest of these speculations are very hard to test. So stay tuned. We'll get back to that shortly.

But first, let's step back and think, OK, if music is an evolved capacity, it should be innate in some sense, at least genetically specified, right, because that's what evolution does is that natural selection acts on the genome to produce things that are genetically specified. And it should be present in all human societies, since the branching out of human societies is very recent in human evolution.

So is it? Well, is music an innate? So, suppose we found specialized machinery in the brain and adults for music. And we showed really definitively, it's really, really, really specialized for music. Would that prove innateness? No, why not?

AUDIENCE: Might have [INAUDIBLE].

NANCY Bingo, thank you, very good. Yup, exactly. So this is something that many, many people are confused about, including colleagues of mine, most of the popular scientific press.

Just because there's a specialized bit of brain that does x doesn't mean x is innate. It could be learned. And the clearest example of that is the visual word form area. Everybody get that?

OK, so we've got to try something else. What if we find sensitivity to music, in some very music particular way, in newborns? Now that will get closer, but here's the problem. Fetuses can hear pretty well in the womb.

And if the mom is singing or even if there's music in the ambient room, some of that sound gets into the womb. So that means that even if you show sensitivity to music, even in some very particular way, in a newborn, it's not a really tight argument that it wasn't, in part, learned.

So this is a real challenge. It may just be impossible to answer. I'm not sure. I don't know how-- I don't know what method could actually answer this. But at the very least, it's really difficult and nobody's nailed it.

So we can backtrack and ask the related, not quite as definitive question: "But OK, how early developing is it?" So often, developmental psychologists take this hedge. It's like, we can't exactly establish definitive innateness. But if things are really there very early and develop very fast, that's a suggestion that at least the system is designed to pick it up quickly. So even if there's a role for experience, there's some things that are picked up really fast and some things that aren't.

And so how quickly is it picked up? So it turns out there's a bunch of studies that have looked at this. And young infants are in fact highly attuned to music. They're sensitive to pitch and to rhythm.

And in one charming study, they took two to three-day-old infants who were sleeping, put EEG electrodes on them, and played them. They wanted to test beat induction, which is when you hear a rhythmic beat. You get trained to the beat. And you know when the next beat is.

And that's true even if it's not just a single pulse. So they played these infants sounds like this. Oh, but the audio is not on. Now it's going to blast everyone. All right, hang on.

AUDIENCE: It's playing.

NANCY Oh, it is playing? Turn up more? OK. Didn't want to deafen people. OK, here.

KANWISHER:

AUDIENCE: It's going a little [INAUDIBLE]. Just turn it up so you can hear it.

AUDIENCE: Go to HDMI, [INAUDIBLE] plugged in [INAUDIBLE].

NANCY It's not, but that's supposed to work, right? It has worked before

KANWISHER:

AUDIENCE: In there.

AUDIENCE: Let's just check your system settings really quickly. So I can hear you from my system.

NANCY Yeah, it's weird.

KANWISHER:

AUDIENCE: Wait, if I can hear you from my system, you're--

NANCY Then, it is going out, yeah.

KANWISHER:

AUDIENCE: Oh, somebody unplugged both. OK, let's try [INAUDIBLE].

NANCY Aah.

KANWISHER:

AUDIENCE: OK, try it one more time.

NANCY OK, here we go.

KANWISHER:

[MUSIC PLAYING]

Did you hear that glitch? Let me do it again. Take it back here.

[MUSIC PLAYING]

Everybody here the hiccup in the beat? So that's what these guys tested. They played rhythms like that to two to three-day-old infants. And--

[MUSIC PLAYING]

Oh, now it's working. OK, great. OK, anyway, so here's what they find with their ERPs. This is the onset of that little hiccup, the time when that beat was supposed to happen and didn't, the missing beat right there. And this is an ERP response happening about 200 milliseconds later for that missing but expected beat. And let's see, this is a standard where the beat keeps going.

Now you might say, well, of course they're different. One has a beat there and one doesn't. They're acoustically different.

So they have a control condition which has a beat, but a different preceding context. So where that beat is not-- I'm sorry, where it has a missing beat, but that's expected by the previous context. So that's just evidence that even young infants have some sense of beat.

So moving a little later, by five to six months, infants can recognize a familiar melody, even if it's shifted in pitch from the version that they learned. And that's really cool, because that means they use relative pitch, not absolute pitch.

And that's something that adults do in music. We're very good at that. But no animal can do that. You can train animals to do various things like recognize a particular pair of sounds or even a few sounds, a few pitches. But if you transpose it, they don't recognize that.

Yeah, Ben.

AUDIENCE: Isn't it possible that we're just sensitive to rhythm and pitch rather than being sensitive to music itself?

NANCY Yes, hang on to that thought. It takes more work to show that it's music per se rather than just rhythm and pitch.

KANWISHER: We'd have to say what we meant by rhythm.

If we load enough into the idea of rhythm, then it's like most of music right there. But we might say just even beat. How about that, right?

And actually, already this study already is not just an even beat, because it has more context than that. That is, for example, the beats in this ERP infant study were not emphasized louder. The infants have to be able to pick out what the beat is from that complex sound.

It's not automatically there in the acoustic signal as the louder onset sound.

Five-month-old infants, if you play them a melody for one or two weeks, so they get really familiar with it and learn it, and then you don't play it again and you come back eight months later, they remember it. So music is really salient to infants.

On the other hand, newborn infants' appreciation of music is not-- what is that not doing there? Oh, yeah, that's right. So they don't prefer consonance over dissonance, right. And they're insensitive to key. And they detect timing changes in rhythms, whether they are timing changes that are typical in the kind of music they've heard or typical in a more foreign kind of music.

And so a really nice study that shows this is that in Western music, it's really common to have-- most Western music has isochronous beat. So you can see that over here. Here's an isochronous beat.

Those are even, temporal intervals. And there's a whole note here and then half notes. And they're all multiples of each other, just wholes and halves, with the beat happening every four notes.

Non-isochronous beat has this funny business where there's a whole note and a half note, making up just three-- what do you call those things-- they're not beats. What are they called?

AUDIENCE: Three-beat notes.

NANCY Sorry, three notes, I guess. But it's not even notes, because it's whatever. I don't know what the terminology is.

KANWISHER: But anyway, this sound here followed by 4.

This is non-isochronous rhythm. Those are really common in Balkan music where they do all kinds of crazy things, like 8/22 or something like that. I mean, like really, really crazy musical meters. They're awesome, I love them.

But they are very other. Like, if you grew up in Western society when you first hear Balkan rhythms, it's very hard to copy them. But six-month-old infants get rhythms equally well if they're isochronous or non-isochronous.

By 12 months, they can only get automatically, like immediately, perceive and appreciate rhythms that are familiar from their cultural exposure. That is isochronous if they're from a Western society or non-isochronous if they're from a Balkan country. Yeah?

AUDIENCE: just what is getting a meter again?

NANCY Well, so there's a whole bunch of studies. I'm just summarizing here. That is, they're sensitive to violations by all

KANWISHER: kinds of measures of little whatever behavioral thing you can get out of a five-month-old, whether it's how much they're kicking their legs or how much-- often, it's how hard they're sucking on a pacifier is another measure.

So you just see, can they detect changes in a stimulus or violations by any of those measures. Or you could do it with the ERPs.

So brief exposure to a previously unfamiliar rhythm is enough for a 12-month-old to appreciate the relevant distinctions in that rhythm, but not for adults.

So if you haven't heard non-isochronous Balkan rhythms until now and you try dancing to them, good luck to you. You can probably get it eventually, but it will take you a long time.

So does this sound familiar? Perceptual narrowing, right? So we keep encountering this. We encountered this with face recognition, with same versus other races, same versus other species. You see it in face recognition.

We encountered it with phoneme perception. The phonemes-- remember, newborn infants can distinguish all the phonemes of the world's languages, even those exotic clicks that I played last time from Southern African languages. And you guys can't distinguish all those clicks now. So that's perceptual narrowing.

It makes sense, of course, because the reason we have perceptual narrowing is you want to have invariants. You want to appreciate the sameness of things across transformations. And if your speech culture or your music culture is telling you these two things, this variation, doesn't count, you want to throw away that difference and treat them as the same. And then once you do that, you can't make that discrimination anymore.

So on this question we started with, is music an evolved capacity. If so, it should be innate. And we haven't really answered that question, maybe. But as I said, it's really hard, and maybe ultimately unanswerable. But certainly it's early developing.

What about this other question? Is it present in all human societies? Well, I said before briefly that it is. Oh yeah, sorry, we have to back up and say, OK, to answer this question, we have to say what is music.

To answer whether it's present in all societies. And this has been a real problem, because music is notoriously hard to define. And many people have made a point of stretching the definition of music, including the ridiculous and hilarious John Cage.

So this is his 1960 TV appearance.

[VIDEO PLAYBACK]

- Over here, Mr. Cage has a tape recording machine, which will provide much of the-- will you touch the machine so we can know where it is-- which will provide much of the background. Also, he works with a stopwatch. The reason he does this is because these sounds are in no sense accidental in their sequence. They each must fall mathematically at a precise point. So he wants to watch as he works.

He takes it seriously. I think it's interesting. If you are amused, you may laugh. If you like it, you may buy the recording. John Cage and "Water Walk."

[EXPERIMENTAL MUSICAL SOUNDS]

[END PLAYBACK]

NANCY Anyway, it goes on and on like that. I guess it was a little edgier in 1959 than it is now. But he's making a point.
KANWISHER: He's making a point is, what the hell is music.

And he's saying, I can call this music if I want. And everybody's enjoying it. Anyway.

So you can watch the YouTube video, if you want. It's quite entertaining. Despite this kind of nihilistic view that anything could count is music, there are some things we can say.

First thing I'd say is, if you want to study music, one of your first things you run into is, oh, what's going to count. You run into this problem here. But actually, I think that doesn't need to be so paralyzing as it feels at first. You can just take the most canonical forms where all of your subjects will agree that this is music and this isn't. And then someday you can study the edge cases later, but you don't need to agonize about them in order to get off the ground and study it.

Further, we can ask what is music cross-culturally. Oh, right, I keep forgetting my next point. And let me make another point is that music is not just about a set of acoustic properties. You may think of music as just an auditory thing, a solitary experience, because a lot of the time it's like that.

But remember that that's a very recent cultural invention. And throughout most of human evolution, music has been a fundamentally social phenomenon, more like this, experienced in groups of people as a kind of deeply social, communicative, interactive kind of enterprise. Or even if not in a large group, music is very social in this sense here.

There's a whole bunch of cool studies about the role of song in infants and how infants use song to glean information about their social environment. And the point is just music is extremely social. It's not just defined by its acoustic properties.

But in addition, we can ask, OK, let's look across the cultures of the world and ask, are there universals of music? Is there anything in common across all the different kinds of music that people experience in different cultures? For example, are there always discrete pitches or always isochronous beats. I already showed you there aren't always isochronous beats.

And this is nice because it's an empirical question. There's a really cool paper from a few years ago where they took recordings of music from all over the world, all those colored dots, and they asked, what are the properties that are present in most of those musics and how prevalent are they.

And what they found is there's no single property of music that's present in all of those cultures, but there's many that are present in most, and there are a lot of regularities. So this is a huge table from their paper where they list many different possible universals.

And what you see is the relevant column is this one here. And the white is the percent of those 304 cultures that they looked at that have that property in their music. So these top ones are very prevalent, just not quite universal, because there's a couple of cases that don't have it.

So one of the most common ones is the idea that melodies are made from a limited set of discrete pitches, seven or fewer, and that those pitches are arranged in some kind of scale with unequal intervals between the notes. So that's as close to a universal of music as you can get, although you can see from that little teeny black snip that it's not quite perfectly universal.

And the second thing is that most music has some kind of regular pulse, either an isochronous beat or even the non-isochronous ones have different subdivisions with different numbers of beats so that there's a systematic rhythmic pattern. So there's something kind of like melody and something kind of like rhythm in almost all the world's musics.

They did find some pretty weird ones, one I can't resist playing for you. This is from Papua New Guinea. So as they say, the closest thing to an absolute universal was song containing discrete pitches, or regular rhythmic patterns, or both, which implied to almost the entire sample.

However, music examples from Papua New Guinea contain combinations of friction blocks, swung slats, ribbon reeds, and moaning voices-- I don't know what those things are either, but I'll play them for you in a second-- that contained neither discrete pitches nor an isochronous beat. OK, here we go.

[VIDEO PLAYBACK]

[PAPUA NEW GUINEAN MUSIC]

[END PLAYBACK]

OK, pretty wild, huh? So maybe wilder, arguably, than John Cage. But anyway, so there are some like pretty remote edges to the concept of music.

I mentioned before the case of consonance and dissonance and that infants don't prefer one over the other. In fact, this links to a really cool recent study from Josh McDermott's lab. And so the question he asked is, why do we like consonant sounds like this-- oops, [INAUDIBLE] play. Here we go.

[RHYTHMIC SOUND]

Kind of nice, right? But we're not so hot about this.

[OFF TUNE SOUND]

Right, everybody get that intuition? OK so what's up with that? So many people have hypothesized for a long time that that difference is based in biology, or even it's like a physical analog of it, beats and stuff like that. But actually, it's an empirical question. And so one way to ask that question is to go to a culture that's had minimal exposure to Western music, all of which really prefers consonance over dissonance. Yes, [? Carly? ?]

AUDIENCE: Is consonants [INAUDIBLE] differentiated [INAUDIBLE]?

NANCY Oh, yeah, yeah. I'm sorry, totally different word-- consonance, C-E, has no relationship to consonants as

KANWISHER: distinguished from vowels. A consonant and a vowel, those are two different kinds of phonemes.

Here, consonance is that difference between those two sounds I just played. And it has to do with the precise intervals of those harmonics in the harmonic stack. All right, so what McDermott and his co-workers did is to go to a Bolivian culture in the rainforest in a very remote location to test these people here, the Tsimane'. And the Tsimane' lack televisions and have very little access to recorded music and radio.

Their village doesn't have electricity or tap water. You can't get there by road and you have to get there by canoe. So that's what McDermott and his team did. They went down there to visit the Tsimane'.

And what they found, they played them consonant sounds and dissonant sounds, and with a translator, and spent a lot of time making sure that they really understood the difference between liking and not liking. And they tested their understanding of what it means to like something or not like it, and all kinds of other ways. And the upshot is, the Tsimane' do not have a preference for consonance over dissonance.

So it's not a cultural universal. And that's consistent with the idea that it's not a preference in infants either. So this is something specific to Western music.

So that's kind of introduction to some stuff about what music is and what its variability is and the fact that its presence is universal. And there are many very common properties across the world's musics, and it developed early.

So let's ask, is music a separate capacity in the mind and brain. All right, so let's start with the classic way this has been asked for many decades, and that's to study patients with brain damage. And it turns out there is such a thing as amusia, the loss of music ability after brain damage.

And so there are both sides of this. There are people who have impaired ability to recognize melodies without impaired speech perception. And there's the opposite-- people who have impaired speech recognition without impaired melody recognition.

So that is, of course, a double dissociation, sort of, it's a little mucky in there. If you state the word simply like that, if you look in detail, there's some muck, as there often is.

So let's look in a little more detail at these two cases, the most interesting ones who seem to have problems with auditory tunes but not with words or other familiar sounds. So here is a horizontal slice. This is an old study. So it's a CAT scan showing you something's up with the anterior temporal lobes in this patient.

And this was true of these two classic patients, CN and GL. Both of them were very bad at recognizing melodies, even highly familiar melodies, happy birthday and stuff like that, they don't recognize. They mostly have intact rhythm perception. And this is a core question we'll come back to. It's a complicated non-resolved situation.

But these guys had intact rhythm perception and relatively intact language and speech perception. However, upon further testing, it becomes clear that these guys have a more general problem with pitch perception, even if it's not in the context of music.

So this is a question that I asked all of you guys to think about for in the opposite direction in your assignment for Sunday night. When I asked you whether those electrodes in the brains of epilepsy patients that are sensitive to speech prosody, to the intonation contour in speech, I asked you whether you thought they would also be sensitive to the intonation contour in melodies. And most of you said, yes, it's pitch, pitch contour, must be.

Well, it's a perfectly reasonable speculation, but not necessarily. Maybe we have special pitch contour processing for speech and different pitch contour processing for music. It's possible. It's an empirical question.

Was there a question back there a second? OK, so maybe this is about pitch for both speech and music, not music per se. And so there are more detailed studies of patients with congenital amusia.

And just like the case with acquired prosopagnosia versus congenital prosopagnosia, whether you get it from brain damage as an adult or whether you just always had it your whole life, and nobody knows exactly why and there's no evidence of any brain damage, the same thing happens with a congenital amusia. So something like 4% of the population, they might say they're tone deaf.

But just to tell you what that means, it can be really quite extreme. They can just completely fail to recognize familiar melodies that anyone else could recognize. They may be unable to detect really obvious wrong notes in a canonical melody. They're just really bad at all of this.

And further, they don't have whopping obvious problems with speech perception. So at first, it was thought that speech perception was fine. But if you look closer, it looks like actually there is, even outside of music, there is a finer grained deficit in pitch contour perception that shows up even in speech.

So what I mentioned before, so we can ask this in the case. This is sort of the reverse case of the ones you considered. Now we have people who have this problem with pitch contour perception in music. Are they going to have a problem also with pitch contour perception in speech? So that's what this study looked at.

So they played sounds like this. And you have to listen carefully. There will be sentences spoken. And you have to see if they're identical or different. So listen carefully.

[VIDEO PLAYBACK]

- She looks like Ann. She looks like Ann?

[END PLAYBACK]

NANCY

KANWISHER:

How many people thought that was different? Good, you got it. So one is the statement and one is-- it's sort of a question. It's in a sort of British accent. It's a little harder to detect, but different intonation contour.

So that's what the Tang, et al. Paper was talking about is that distinction. So we can then ask, that subtle distinction, are people with congenital amusia impaired at that. So if it's specific to music, they shouldn't be. But if it's any intonation contour, they should be.

Yeah, I'll play the other ones. So they are in fact impaired. This is accuracy here, the controls are way up there, the amusics are down there.

So they are impaired at this pitch contour perception thing, even in the context of music. I'm sorry, I said that wrong-- even in the context of speech. So it's not just about music.

And in the controls, they have sounds like this, which are just tones. Got that? It's the same kind of thing, but not speech. And you see a similar deficit in the amusics compared to the controls. And then they have a nonsense speech version.

[VIDEO PLAYBACK]

- [INAUDIBLE]

[END PLAYBACK]

NANCY Same deal-- the amusics are impaired compared to the controls. So that shows that the deficit for these guys is

KANWISHER: not specific to music per se but it seems to be a pitch contour problem in general that extends to speech. Yeah?

AUDIENCE: Which of those--

NANCY We'll get there, sort of. It would have been nice if the Tang et. al. paper had included some musical contour stuff.

KANWISHER: They didn't, but I'll show you some of our data shortly that gets close to this.

OK, so all of that suggests that this amusia is really more about pitch than speech. I'm sorry, what's the matter with me. It's really more about pitch than music.

But the reading that I assigned for today is a very new twist in this evolving story. So this used to be a nice, clean lecture with a simple conclusion. And now all of a sudden, I ran across that paper. It's like, wow, OK, that might not be quite the case.

So what did you guys get from the reading? In what way does that slightly complicate the story here? Yeah, [INAUDIBLE]?

AUDIENCE: [INAUDIBLE]

NANCY Yeah, what they found is that amusics, not all of them, also have problems with rhythm. And that is inconsistent

KANWISHER: with the idea that amusia is just about pitch, whether in speech or music. And that says, OK, many amusics also have problems with rhythm. Yeah?

AUDIENCE: [INAUDIBLE]

NANCY So there's a standard battery that people use that asks-- Dana, help me. What does the standard battery ask

KANWISHER: people?

AUDIENCE: There's a lot stuff, tests, things like listening to like a clip of a symphony and having to decide whether [INAUDIBLE] or they're too slow.

NANCY Kinds of things that people without musical training answer fine, although there's quite a range. I'm at the way

KANWISHER: bottom end of Dana's scale when she gives these.

AUDIENCE: That rhythm falls apart, might not be able to tell the difference.

NANCY Just that this prior evidence on the stuff I showed and a whole bunch of other studies seem to suggest that

KANWISHER: amusia, both in acquired brain damage and congenital amusia, seem to be really when you drill down more of a problem with pitch per se, even pitch in speech.

And so then if it's about pitch, why would it also go along with rhythm? And so when it goes along with rhythm, that starts to sound more like this is something about music. It gums up the story. Talia?

AUDIENCE: So I don't really know if this could be a compound, but when it comes to natural speech when you have some kind of intonation, like pitch differences when you emphasize, like especially in terms of a question, aren't there also some kind of rhythmic differences as well?

NANCY Yeah.

KANWISHER:

AUDIENCE: So how do you separate the two out?

NANCY You just have to do a lot of work to try to separate those out. And so the paper I signed to you guys did some of that work. There's still room to quibble, but they did. There was experiment two, and they tried to deal with exactly that kind of thing of saying, OK, let's try to make sure that-- well, actually the controls that they were doing is slightly different. They were to make sure that the beat task didn't require pitch. So it's very, very tricky to pull these things apart, which is--

AUDIENCE: Yes, so like the beat task doesn't make sense, but I was just, like, in the verb first one, even from the paper that was assignment Sunday. I don't know, so you're saying that it's totally possible to separate out rhythmic differences from when you're just changing pitch.

NANCY It's really, really difficult. It's really difficult. Dana's trying to do experiments to do this right now. And she's invented some delightful and crazy stimuli that try to have one and not the other. It's very tricky.

KANWISHER:

You can have rhythm without pitch change. That you can totally do. It's really hard or impossible to have a melodic contour without some beat or other. We have some crazy stimuli that sort of do that, but they're pretty crazy.

So anyway, these are very tricky things to pull apart. And this is all right at the cutting edge. These things have not been cleanly separated.

I'm running out of time. So do you have a quick question? OK, sorry about that.

So conclusions from the patient literature, they're suggestive evidence for specialization for music, but no really clear disassociations. Music deficits are frequently but not always associated with just more general pitch deficits.

And all of this is complicated because there's lots of possible components of music, right. When there's pitch deficits, is it pitch or relative pitch, interval, key, melody, beat, meter? All of these things are different facets of music.

And so it's really not resolved exactly what's going on here. It's kind of encouraging that there's a space in there, but not resolved.

So let's go on to functional MRI. And we're going to run out of time. So let me just take a moment to figure out how I'm going to do this.

What the hell am I going to do here? Well, I hate to-- OK, you guys are going to tell me at 12:05. Yeah, OK. Maybe we can get all through this.

So here's a really charming study from a few years ago that tried to ask whether there are systematic brain regions that are engaged in processing music. And they used a really fun perceptual illusion that you're going to hear. I'm going to play a speech clip.

And it's part of it is going to be repeated many times. And just listen to it and think about what it sounds like.

[VIDEO PLAYBACK]

- For it had never been his good luck to own and eat one. There was a cold drizzle of rain. The atmosphere was murky. There was a cold drizzle. There was a cold drizzle.

[END PLAYBACK]

NANCY What happened?

KANWISHER:

AUDIENCE: [INAUDIBLE]

NANCY Yeah? What happened?

KANWISHER:

AUDIENCE: [INAUDIBLE]

NANCY You start to hear a melody. And you didn't hear the melody the first time he said it. It was just normal speech,
KANWISHER: right. Speech has this kind of intonation contour. And he's speaking with an intonation contour. But then somehow when you keep hearing it, it turns into a melody.

So it turns out that doesn't work for all speech clips. In fact, it's really hard to find speech clips for which it works. But there are some. But everyone has that experience, or most people do. And that gives us a really nice lever, because we can take that same acoustic sound when you hear it as speech and when you hear it as melody and we can ask, are there brain regions that respond differentially.

It's sort of analogous to upright versus inverted faces. Well, it's even better. It's the exact same sound clip that's construed one way at first and another way afterwards. Everybody get that?

So that's what these guys did. They used a standard block design. They just listened to those sounds and they just looked in the brain to see what bits respond more after the sound starts getting perceived as music than before when it was being heard as speech.

And they got a bunch of blobs in the brain. It's a bit of a mess, but they got some stuff. And so that's fun.

But it's also ambiguous. We still don't know if this is about some kind of pitch processing, which becomes more salient-- you hear it as abstract pitch-- or whether it's really about melodic contour or what. So that's a cool study, but I think it doesn't really nail what's going on.

So another angle at this is to ask whether music recruits neural machinery for language. So let me say why this has been such a pervasive question in the field. So there's a lot of people who have pointed out for 30 years, or probably more, there are many deep commonalities between language and music.

So they're both distinctively or uniquely human. They're natively auditory. That is, we can read language, but that's very recent. Really, language is all about hearing, evolutionarily. They unfold over time.

And they have complex hierarchical structure. So you can parse a sentence in various ways and there are all kinds of people who've come up with ways to have hierarchical parsings of pieces of music as well. So there's a lot of deep connections between language and music. And so many people have hypothesized that they use common brain machinery.

And there, in fact, many reports from neuroimaging that argue that in fact they do use common machinery. Like, we found overlapping activation in Broca's area for people listening to music and speech.

However, both studies are all group analyses. I forget if I've gone on my tirade in here about group analyses. Have I done the group analysis tirade in here?

You'll get more of it later. I'll do a brief version now, and you'll get more later.

Here's the problem-- group analysis is you scan 12 subjects. You align their brains as best you can. And you do an analysis that goes across them. And you find some blob, say, here, yeah, be there, for listening to sentences versus listening to non-word strings. OK, that's a standard finding.

Then you do it again for listening to melodies versus listening to scrambled melodies. And you find the blob overlaps. And then you say, hey, common neural machinery for sentence understanding and for music perception.

Now that's an interesting question to ask. It's close to the right way to do it. but there's a fundamental problem. And that is, you can find an overlap in a group analysis, even if no single subject shows that overlap at all. Why? Because those regions vary in their exact location. And if you mush across a whole bunch of individuals, you're essentially blurring your activation pattern.

And so all of the prior studies, until a few years ago, had been group analyses and they found overlap. And who the hell knows if there was actually overlapping activation within individual subjects, which there would have to be if it's common machinery. Or if they're just nearby and you muck them up with a group analysis and they look like they're on top of each other.

If you didn't quite get that, we'll be coming back to that point. For now, all you need to know is many people ask this question and the methods were close but problematic. But luckily, however, Ev Fedorenko did this experiment right a few years ago.

So here's Ev and here's what she did, she functionally identified language regions in each subject individually. And we'll talk more about exactly how you do that. You listen to sentences versus non-word strings. You find a systematic set of brain regions that you can identify in each individual that look like this.

Here is in three subjects. Those red bits are the bits that respond more when you listen to a sentence versus listen to non-word strings or read sentences versus non-word strings.

Then what she could do is she said, now that I found those exact regions in each subject, I can ask of those exact regions, how do they respond to music versus scrambled music. So she played stuff like this.

[MUSIC PLAYING]

OK, so nice canonical and nothing crazy, weird. We're not going with the New Guinean music and asking edgy questions. We're just saying something everybody agrees that's music, versus you scramble it and it sounds like this.

[MUSIC PLAYING]

OK, it's actually the same notes. I know, I know. A lot of people that go, that's cool, that's really edgy. Yeah, it is. But to most people, it's not canonical music.

And so what Ev found is that none of those language regions responded more to the intact than scrambled music. So language regions are not interested in music. We'll talk more about that next week or the week after.

Then she did the opposite. She identified brain regions here in a group analysis just to show you where they are, anterior in the temporal lobes, that respond more to intact than scrambled music.

She identified those in each subject and measured the response of those regions to language, sentences and non-word strings. And each of those regions respond exactly the same to sentences and non-word strings.

So basically, the language regions are not interested in music, and the music regions are not interested in language. And therein, we have a--

AUDIENCE: [INAUDIBLE]

NANCY
KANWISHER: Thank you, exactly. So music is not using machinery for language. That was one of the hypotheses we started with. And it was not.

So that's true, at least for high-level language processing, that computes the meaning of a sentence. But what about speech perception? Remember, last time I made the distinction between the sounds, like ba and pa, which have a whole set of computational challenges, just perceiving those sounds, which is quite different than knowing the meaning of a sentence.

So what about speech perception or, in fact, any other aspect of hearing? So what I'm going to try to do is briefly tell you about one of our experiments. I'm sorry, I try not to turn this whole course into stuff we've done in my lab, but it's one of my favorite ever. And it's a cool, different way to go at this question from the other MRI experiments we've talked about before.

So the background is, OK, let's step back. What's the overall organization of auditory cortex? And when we did this experiment five or six years ago, not a whole lot was known. Basically, everybody agrees. Whoops, I put the wrong slide in here. Everybody agrees that primary auditory cortex is right there with that high-low-high frequency thing we talked about from there.

But from there on out, in the last couple of years, there's an agreement about speech selective cortex that I showed you briefly last time and other people have seen that. But there's lots of hypotheses and no agreement with anything else and no real evidence for really music-selective cortex.

But there's a problem with all the prior work where you sit around and make a hypothesis and say, oh, let's see, are we going to get a higher response to, say, intact versus scrambled music, or faces versus objects, or whatever. All of those are scientists making up hypotheses, and then testing them.

And there's nothing wrong with that. That's what scientists are supposed to do-- invent hypotheses, and then make good designs and go test them. But the problem with that is, we can only discover things that we can think to test.

What if deep facts about mind and brain are things that nobody would think up in the first place? And so that's where we can get real power from what are known as data-driven studies, where you collect a boatload of data and then use some fancy math and say, tell me what the structure is in this data.

Not, is this hypothesis that I love true in these data. And I'll do anything to pull it out if I can. See it in there, find evidence for it in there. But yeah, exactly.

But if we collect a whole bunch of data and do some math and see what the structure is, what do we see? So that's what we did in this study. I'm going to speed up to try to give you the gist here.

So "we" is Sam Norman-Haignere here and Josh McDermott.

[SOUND RECORDING EXPERIMENT PLAYING]

And so we scanned people while they were hearing stuff like this. We first collected the 165 categories of sounds that people hear most commonly. This is classic cocktail party effect you guys are doing. You have to separate me speaking from all this crazy, weird, changing background.

And so anyway, we scan people listening to these sounds, which broadly sample auditory experience. And so we collected sounds people hear most often and that they can recognize from a two-second clip. OK, enough already.

[CELLPHONE RINGING]

Oh, yeah, just to wake everyone up. So we scan them listening to those 165 sounds, broad sample of auditory experience. Then, from each voxel in the brain, we measure the exact magnitude of response of that voxel to each of the 165 sounds and we get a vector like this. Everybody with me? That's one voxel right there, another voxel, another voxel.

We do this in all of kind of greater, suburban, auditory cortex. That is not just primary cortex, but all this stuff around it that might even remotely, that responds in any systematic way to auditory stimuli. They grabbed the whole damn thing.

So you do that in 10 subjects. You have a big matrix like this-- 1,000 voxels in each subject, 11,000 voxels across the top, 165 sounds. That's our data.

So each column is the response of one voxel in one person's brain to each of the 165 sounds. Everybody got it?

Now, we have this lovely matrix, which is basically all the data we care about from this whole experiment. Then, we throw away all the labels. Poof. It's just a matrix.

And then we do some math, which essentially says, let's boil down the structure in this matrix and discover its fundamental components. That math happens to be a variant of independent component analysis, if that means anything to you. If it doesn't, don't worry about it.

The gist is, we're doing math to say what's the structure in here. And we're doing it without any labels. So this analysis doesn't even know where the voxels are or which of your 10 subjects that voxel came from. It doesn't know which sound is which.

And so it's very hypothesis neutral. It's a way to say, show me structure with almost no kind of prior biases. Just show me the structure. So everybody get how that's kind of a totally different thing to do from everything we've talked about so far?

So that's what we did. I'm going to skip the math and the modeling assumption. It's not really that complicated, but I think I'm going to run out of time, so very hypothesis neutral.

And what we find is six components account for most of the replicable variance in that whole matrix. I'll tell you what a component is in a second. Did you have a question?

AUDIENCE: Is it just like with ICA, but [INAUDIBLE] PCA [INAUDIBLE]?

NANCY With PCA, you assume orthogonal axes. With ICA, you don't assume orthogonal axes. And so it's very, very
KANWISHER: similar to PCA. And it starts out as PCA and then it does some more rigmarole. Yeah, it's the same idea. Like basically, tell me the main dimensions of variation. Yeah?

AUDIENCE: And are these matrices sparse and [INAUDIBLE]?

NANCY Yes, they are sparse. And that is one of the assumptions you use. There isn't only one way to factorize a matrix.

KANWISHER: It's an ill-posed problem.

So you need to make some assumptions. And that's one of the ones we made, but you can test them. So what we find is six components account for most of the data.

And four of those reflected acoustic properties of the stimuli. One was high for all the sounds with lots of low frequencies. Another was high for all the sounds with high frequencies. What is that? Sorry, speak up?

AUDIENCE: [INAUDIBLE]

NANCY They're sensitive to frequency, but where is that in the brain that you've already heard about?

KANWISHER:

AUDIENCE: Primary--

NANCY Primary auditory cortex as a tonotopic map. So this is awesome. Because if you go invent some crazy math and
KANWISHER: you apply it to your data and you discover something you know to be true, that's very reassuring. The math isn't just inventing crazy stuff. It's discovering stuff we already know to be true.

That's known in more biological parts of the field as a positive control. Invent a new method, make sure it can discover the stuff you know to be true. So check, check, OK.

But then it discovered some other stuff. And I'm just going to tell you about two of them. So here's one. So I was just loose about what a component is. A component is a magnitude of response for each of the 165 sounds and a separate distribution in the brain, which I'll show you in a moment. So here's one of those components.

And we've taken the 165 sounds and added basic category labels on them. We put them on Mechanical Turk and people told us which category they belong to. So that enables us to look at this mysterious thing and average within a category.

So this is its component. And if you look at it, you see that it's really high for English speech and foreign speech that our subjects don't understand. And then, oh, what's, that intermediate Thing Oh, that's music with vocals.

It has a kind of speech. And way down here-- that's non speech vocalizations, stuff like laughing and crying and sighing. So there's a voice but no speech content. So that's a speech component.

And as I mentioned, this had been seen before in the last few years. So it wasn't completely new. But what's cool about this is just emerged spontaneously from this very broad screen. We didn't go and say, hey, can we find a speech selective region of cortex, if we try really hard. Oh, yeah, we validate our hypothesis.

This is like, let's sample auditory experience-- and wow, there it is. Yeah?

AUDIENCE: I mean, you assigned [INAUDIBLE].

NANCY We put them on Turk and had people say what category they fit into. Yeah?

KANWISHER:

AUDIENCE: [INAUDIBLE]. Categorizing by speech is a very good way [INAUDIBLE] better way than [INAUDIBLE].

NANCY Absolutely, absolutely. This is a first pass. And one hopes to go deeper and deeper. If we could separate different

KANWISHER: aspects of speech, consonants and vowels, fricatives, whatever, there could be much more to be done. Yeah, I got to-- oh, boy, OK.

And when do I have to give them the quiz? It's shortish. They don't need a full 10 minutes. What is it? Seven questions?

AUDIENCE: Eight.

NANCY Eight-- eight minutes?

KANWISHER:

AUDIENCE: [INAUDIBLE]

NANCY OK, make me stop definitively at 12:18. OK, so that's cool. It's not exactly new, but it's a really nice way to

KANWISHER: rediscover things that we thought to be true.

All right, then there's component 6 that popped out. What is component 6? Well, if we average within a category instrumental music and music with vocals, and everything else is really low. We didn't go looking for this. Boom-- music selectivity.

That's pretty amazing. Never really been seen before. People have looked and they've made some kind of sort of smoke and mirrors, like, not really. This is the first time it was seen and it just popped out of the data.

And that says that it's not just something you can find if you try really hard and go fishing for it. It's actually a significant part of the variance in this whole response. I'm going to skip everything except clarification questions now, because I'm-- go ahead.

AUDIENCE: Did these voxels correspond to the music [INAUDIBLE]?

NANCY Sort of, it's complicated. Sorry, it's a long answer. So this really looks like it's music. And so now, I was vague
KANWISHER: about what a component is, but it's both that response profile and it's a set of weights in the brain.

So if you project this one back in the brain, you get this band of speech selective cortex right below primary auditory cortex, like that. And if you project the music stuff back in the brain, you get a patch. This is sort of an answer to your question. You get a patch up in front of primary auditory cortex and a patch behind.

So here we have a double dissociation of speech selectivity and music selectivity in the brain, OK? So music doesn't just use mechanisms for speech as many people have proposed. It's not true, right.

So when you see dramatic data like this, a natural reaction is to say, like, really, get out, come on. Like, music specificity, like what?

So very briefly, Dana has just replicated this in a new sample of subjects. It does not matter if those subjects have musical training, like students from Berklee School who spend like six hours a day practicing, versus people who have essentially zero music lessons ever in their life, you get those components in both groups, maybe slightly stronger in the trained musicians. We're not quite sure yet. But in any case, it is totally present in people with zero musical training.

That doesn't mean it's innate, because people without musical training have musical experience but no explicit training. Skip all of this. Here is her replication. Boom, boom. It's there with and without training.

I'm going to skip all this. You can read it on the slides, if I lost you in here, because I want to show you something else. That music selectivity was not evident if you just do a direct contrast in the same data.

Take all the music conditions, all the non-music conditions, you get a blurry mess. It's not strong. You have to do the math to siphon it off.

And that's OK. But I like to see things in the raw data. And so probably what that means is that the music is overlapping with other things in the brain. And so the direct contrast doesn't work well, the math can pull them apart. But wouldn't it be nice to see them separately?

And so we've been doing intracranial recordings from patients with electrodes in their brain. And I'll just show you a few very cool responses. So this is a single electrode in a single patient. These are the 165 sounds, same ones. This is the time course.

And this is a speech selective electrode. It responds to native and foreign music. Those are the two green ones-- I'm sorry, native and foreign speech. And it responds to music with vocals in pink.

Everybody see how that's a speech selective electrode? So there's loads of those. But we also found these. Here is a single electrode. Look, each row is a single stimulus. Here's a histogram of responses to all the music with vocals, music without vocals, much stronger than to anything else.

You might be saying, well, what about those things. Let's look at what those things are. Oh, even the violations aren't really violations. Whistling, humming, computer jingle, ringtone-- those are sort of music.

So that is an extremely music-selective individual electrode in a single subject's brain. No fancy math that might have invented it somehow. It's just there right in the raw data. Further, and here's the time course, you can see the time course of music with instruments, music with vocals, everything else. Really selective.

So this is the strongest evidence yet for music specificity in the human brain. But there's one more cool thing that came out of this analysis. And that is we found some electrodes that are not just selected for music, but selected for vocal music, selected for song.

And that's really amazing. Because as I started off at the beginning, many people have said that song is a kind of native form of music. The first one to evolve and all that kind of stuff.

And so we did all the controls. It's not the low-level stuff. And there's lots of open questions. We started with this puzzle of how did music evolve, if it did evolve. And we made a little bit of progress.

It doesn't share music machinery with speech and language. If it's auditory cheesecake, as Pinker said, it's auditory cheesecake that not only uses machinery that evolved for something else, but changes it throughout development and makes it very selective.

These guys speculated that song is special. Maybe it is. And sexual selection, who knows? We have no data.