Lecture 16. Music

<u>Outline for Today:</u>

O. Brief review of Last Class on Audition and Speech

I. Music

Evolutionary Puzzle: why do humans make music?

Is it even an evolved capacity?

Is it innate?

Is it universal? *What* is universal?

II. Is music a distinct capacity in mind and brain?

Amusia in patients with brain damage

Congenital amusia

III. Cortical Specializations for Music? Distinct from speech and language?

fMRI

ECoG

IV. Quiz

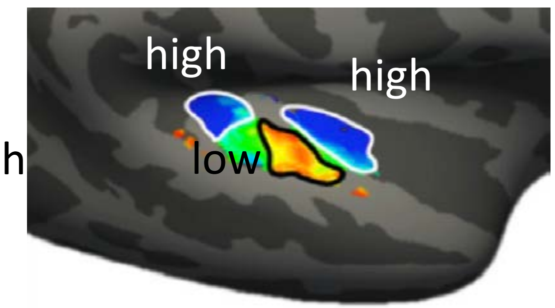
Review of Lecture 15

I. Introduction (computational theory)

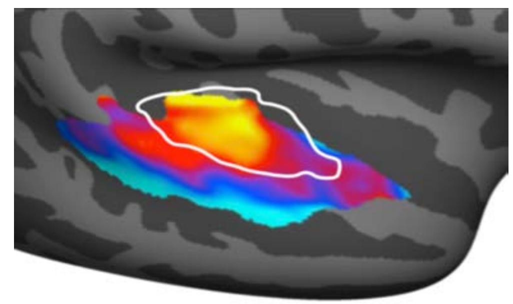
Sound is pressure waves travelling through air
We extract LOTS of info from sound: object and material recognition, locations of sources, etc
These problems are ill-posed, e.g.:
  separating sound sources (cocktail party problem)
  reverb: echoes are on top of original sound, need to pull apart
Both problems solved by using knowledge of the properties of real-world sound.

II. Speech

Phonemes: speech sounds that distinguish words in a given language
Includes vowels (with lots of harmonics)
  and consonants (few harmonics)
Computational challenges:
  talker variability
  interdependence of voice and speech

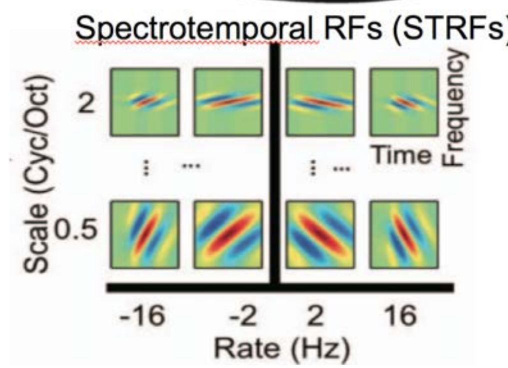


correlation of resp to natural sounds & their STRF-model versions



III. Up to cortex

Primary auditory cortex is:
Tonotopic (high, low, high)
Well modeled by spectrotemporal filter (STRF) model



Today: MUSIC!
An important transition in the course to *uniquely human functions:*
  the coolest things to study (who we are as humans!),
  and the hardest (why?)
Why bother with such a fluffy topic as music?

Why Music?

Not fluffy, but *fundamental*:

Music is a **uniquely** and **universally** human capacity:

- present in some form in every human society
- differs substantially from its closest analogues in animals

Music is important to humans:

- we have been doing it for a long time

  40,000 year old flutes

  singing probably goes back much farther

  maybe even before language???
- arises early in development

  young infants very interested in music

  sensitive to beat and melody (independent of absolute pitch)
- people pay a lot of $ for it

  $43 billion in sales in 2018

These facts raise an obvious question:

Why do humans create and like music?

LOST AND FOUND Scientists say that this bone flute, found at Hohle Fels Cave in Germany, is at least 42,000 years old. Jensen/University of Tubingen

Major puzzle: Why do we have music?

Is music an evolved capacity, specifically shaped by natural selection?
(and if so, what was its selected-for function?)

Many thinkers have struggled with this question, including Darwin

"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…."

Darwin's speculation…

Major puzzle: Why do we have music?

Is music an evolved capacity, specifically shaped by natural selection?
(and if so, what was its selected-for function?)

Yes:

Darwin: sexual selection: *"[I]t 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"*

Mehr & Krasnow (2017): managing parent-offspring conflict:
"*infant-directed song arose in an evolutionary arms race between parents and infants, stemming from the dynamics of parent-offspring conflict*" (infant needs to know parent is attending, parent has other needs)

Et cetera…

No:

Pinker (1994): 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"*

Put another way:
Music is not an evolutionary adaptation at all,
but an alternate use of neural machinery that evolved to serve other functions, like speech and language.

aha! an empirical question!
stay tuned...

If music is an evolved capacity it should be a) innate & b) in all human societies.
Is it?

Music Perception in Infants

Is music innate?

If specialized brain machinery for music in adults would that prove innateness?

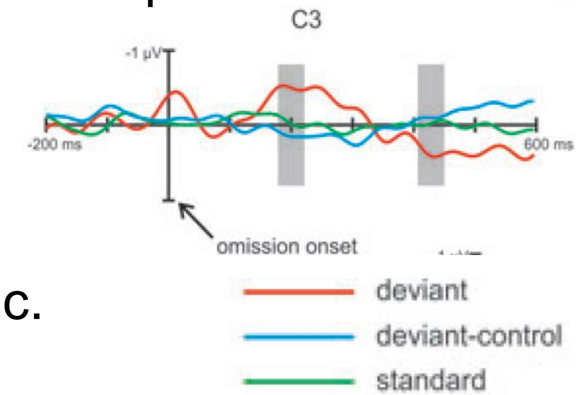
think: VWFA

If we find sensitivity to music in newborns?

problem: fetuses can hear in the womb.

So, a real challenge to answer. (Maybe impossible?)

But we can ask *how early* infants are sensitive to music.



Young infants are highly attuned to music:

are sensitive to pitch and rhythm

2-3 day old sleeping infants show beat induction (ERPs)

by 5-6 months can recognize a familiar melody when it is shifted in pitch

i.e., they use relative pitch, like adults, and unlike animals

5-month-olds familiarized with a melody recognize it 8 months later

Newborn infants' appreciation of music is not culturally specific

infants do not prefer consonance over dissonance

they are insensitive to key

they detect timing changes as well in complex foreign rhythms……

If music is an evolved capacity it should be a) innate & b) in all human societies. Is it?

Infants' Response to Music: Meter

6-month-old US infants "get" rhythmic meters
from unfamiliar nonisochronous rhythm
By 12 months they can only "get" their own culture's rhythms
Brief exposure to unfamiliar meters is
sufficient for 12-month-olds to perceive the relevant distinctions
but not for adults
Sound familiar?
perceptual narrowing!
(same deal with speech phonemes)

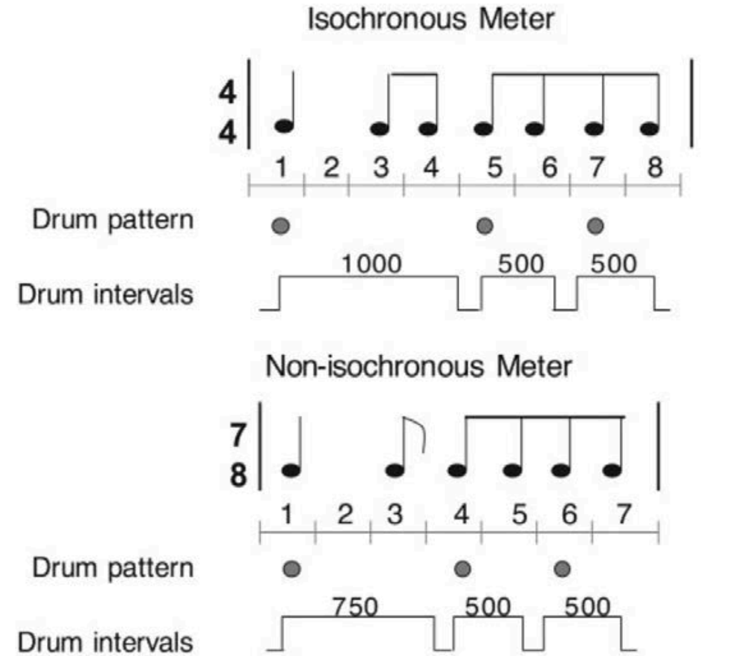


Fig. 1. One measure each of isochronous and nonisochronous meter familiarization excerpts with a long-short-short drum accompaniment, depicted in musical notation and graphical form. Each count of the measure is numbered to illustrate that isochronous meter excerpts consist of eight counts per measure, whereas nonisochronous meter excerpts contain seven counts per measure. The intervals in the isochronous meter drum pattern form a long-to-short ratio of 1,000:500 or 2:1; the intervals in the nonisochronous meter drum pattern form a long-to-short ratio of 750:500 or 3:2.

maybe ???

If music is an evolved capacity it should be a) innate & b) in all human societies. Is it?

What is Music, Anyway?

Notoriously hard to define it…….

John Cage (1959)

Video of John Cage's "Water Walk" as performed in 1960 by the composer on the TV show *I've Got a Secret* © CBS, Inc

But there are some things we can say…..

What is Music All About?

Music is fundamentally *social!*

A reminder……

You may think of music as a solitary enterprise…..

But this is a very recent cultural invention

Throughout most of human evolution,

music has been a fundamentally social phenomenon,

more like this: or this:

discrete pitches?

isochronous beats?

An empirical question!

Are there "universals" of music?

Do Universals of Music Exist?

<u>Savage et al (2015):</u>

304 recordings of music from all over the world:

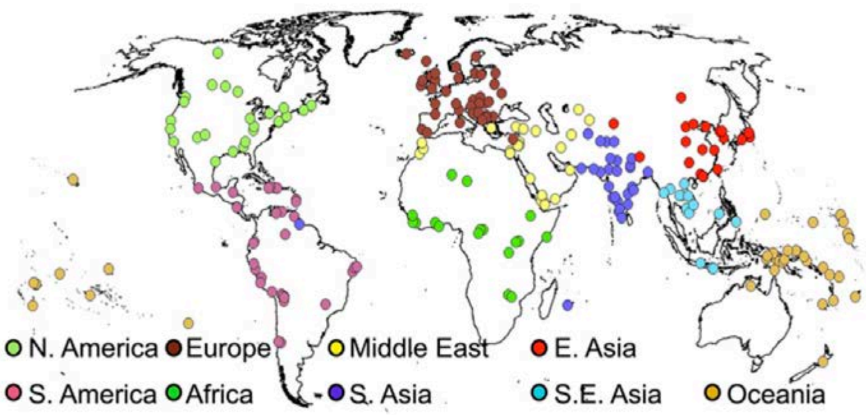
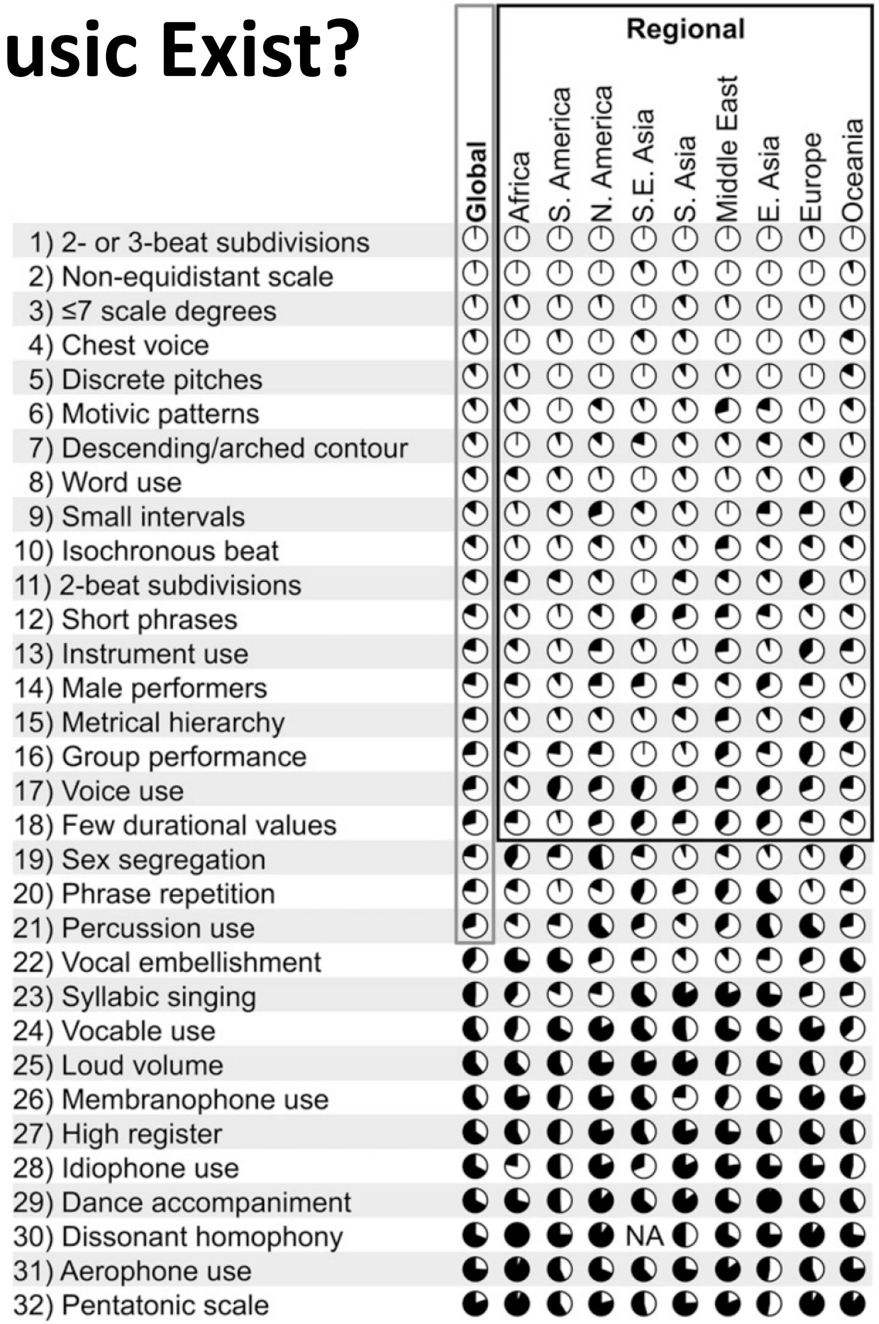


Fig. 1. The 304 recordings from the *Garland Encyclopedia of World Music* show a widespread geographic distribution. They are grouped into nine

Found:

- No absolute universals, but many regularities:

1. Melodies are usually made up of a limited set of discrete pitches (seven or fewer),which form part of a scale that is divided into unequal and relatively small intervals (a perfect 5th or less).
2. Most music also has a regular pulse (an isochronous beat), usually with 2 or 3 subdivisions, and a limited set of rhythmic patterns.



Do Universals of Music Exist?

Savage et al (2015):

304 recordings of music from all over the world:

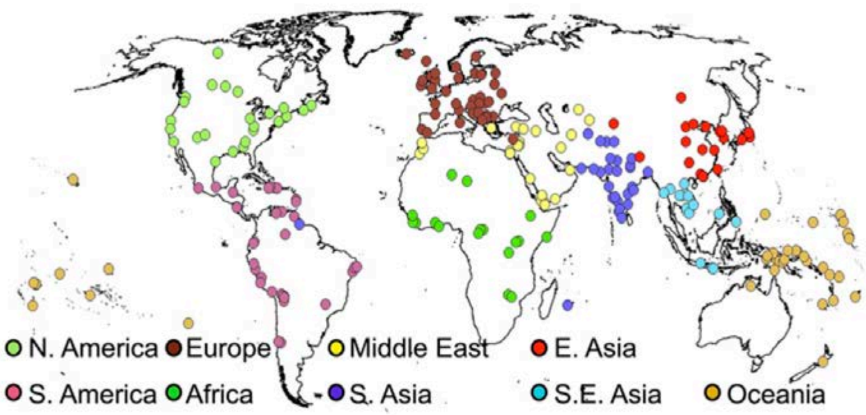
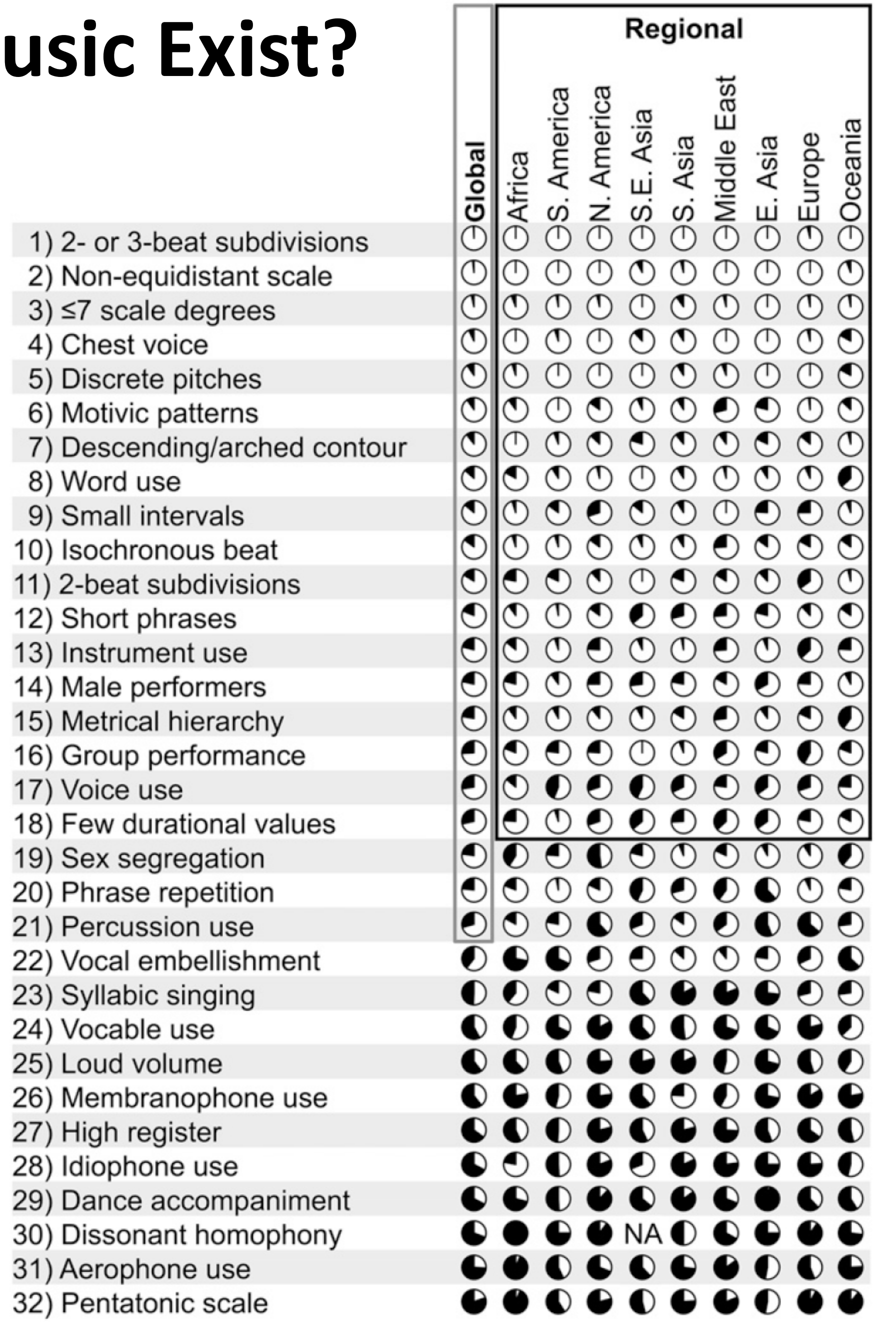


Fig. 1. The 304 recordings from the *Garland Encyclopedia of World Music* show a widespread geographic distribution. They are grouped into nine



- No absolute universals, but many regularities:

"*The closest thing to an absolute universal was [song containing]… discrete pitches or regular rhythmic patterns or both, which applied to almost the entire sample, including instrumental music. However, three musical examples from Papua New Guinea containing combinations of friction blocks, swung slats, ribbon reeds, and moaning voices contained neither discrete pitches nor an isochronous beat.*"

Do Universals of Music Exist?

Test case: Consonance vs dissonance.
Why do we like this (consonant): More than this (dissonant):

Is this preference shaped by biology?
culture (Western music)?

To find out, test the Tsimane: a native society living in a remote village in the Amazon rainforest.

Tsimane' lack televisions, and have limited access to music via radio. Village lacks electricity and tap water, is inaccessible by road, and can be reached only by canoe.

Preference for consonance> dissonance is *completely absent in the Tsimane!*

Consistent with lack of preference in infants.

McDermott et al, Nature (2016)

Lecture 16. Music

<u>Outline for Today:</u>

O. Brief review of Last Class on Audition and Speech

I. Music

- Evolutionary Puzzle: why do humans make music?
- Is it even an evolved capacity?
    - Is it innate?
    - Is it universal? *What* is universal?

II. Is music a distinct capacity in mind and brain?

- Amusia in patients with brain damage
- Congenital amusia

III. Cortical Specializations for Music? Distinct from speech and language?

- fMRI
- ECoG

IV. Quiz

Patient Studies: Acquired ‘Amusia’

Impaired melody recognition without impaired speech perception

Impaired speech recognition without impaired melody recognition

Table 1 Case reports of selective impairment and selective sparing in the auditory recognition of words, tunes and other meaningful sounds

Reports	Auditory domains		
	Tunes	Words	Other familiar sounds
C.N. and G.L.[10]	–	+	+ (+ voices)
I.R.[11]	–	+	+
H.V.[12]	–	+	+
H.J.[13]			
1 case[14]			
K.B.[15]			
11 cases of congenital amusia[17]	–	+	+ (+ voices)
1 case[18]	+	–	+
1 case, during recovery[19]	+	–	+
N.S.[20]	+	–	+
G.L.[21]	+	–	–
1 case[22]	+	–	–
1 case[23]	+	–	–

+ normal; – impaired.

Double dissociation (sort of)

Patient Studies: Acquired ‘Amusia’

What about Congenital Amusia?

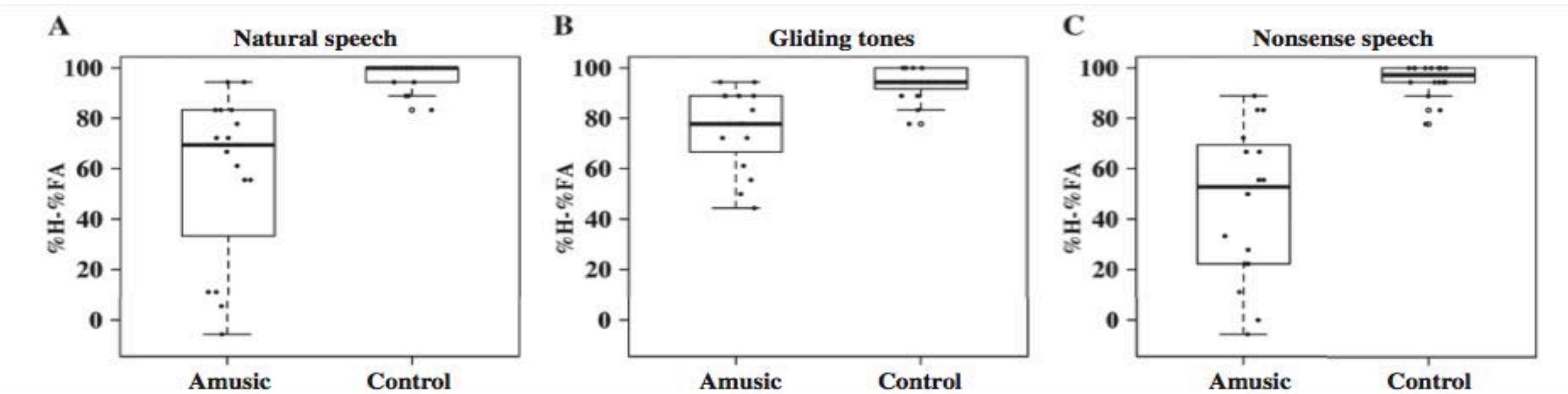
- CN & GL
  - Bad melody recognition
  - Intact rhythm perception
  - Relatively intact language
- But…
  - Probably impaired pitch perception
  - Difficulty with prosodic voice tasks
  - So this may be about pitch (for both music and speech), not music per se

Peretz, 1993; Patel et al., 1998; Dalla Bella & Peretz, 1999

Congenital Amusia

- 4% of population. Inability to recognize familiar melodies, "wrong notes"
- Primary difficulties with music not speech
- But seems to be caused by 'fine-grained' pitch contour deficit
- Most of you all assumed pitch contour in speech = pitch contour in music

True? Measure performance on same/different task on small but ecologically valid intonational pitch contrasts (statement vs question?)



Source: F. Liu, et al. Brain, Volume 133, Issue 6, June 2010, 1682–1693, https://doi.org/10.1093/brain/awq089

Liu et al., 2010

<u>Conclusion</u>: Like acquired amusia, congenital amusia seems to be not a domain-specific deficit in music, but a more general deficit in pitch perception.

But: your reading from today: is it really just pitch?

Conclusions from Patient Literature

- Suggestive evidence for specialization but no clear dissociations
- Musical deficits are frequently associated with more basic difficulties in pitch perception.
- Many possible components of music…
  - pitch, interval, key, melody, beat, meter…..

what can fMRI tell us?

Lecture 16. Music

<u>Outline for Today:</u>

O. Brief review of Last Class on Audition and Speech

I. Music

- Evolutionary Puzzle: why do humans make music?
- Is it even an evolved capacity?
  - Is it innate?
  - Is it universal? *What* is universal?

II. Is music a distinct capacity in mind and brain?

- Amusia in patients with brain damage
- Congenital amusia

III. Cortical Specializations for Music? Distinct from speech and language?

- fMRI
- ECoG

IV. Quiz

Tierney et al (2013): A Musical Illusion

You will hear a speech clip, then a subset of it will be repeated many times, then you will hear the original clip again. Listen carefully:

fMRI Blocked design, just listen and note if it sounds like speech or music

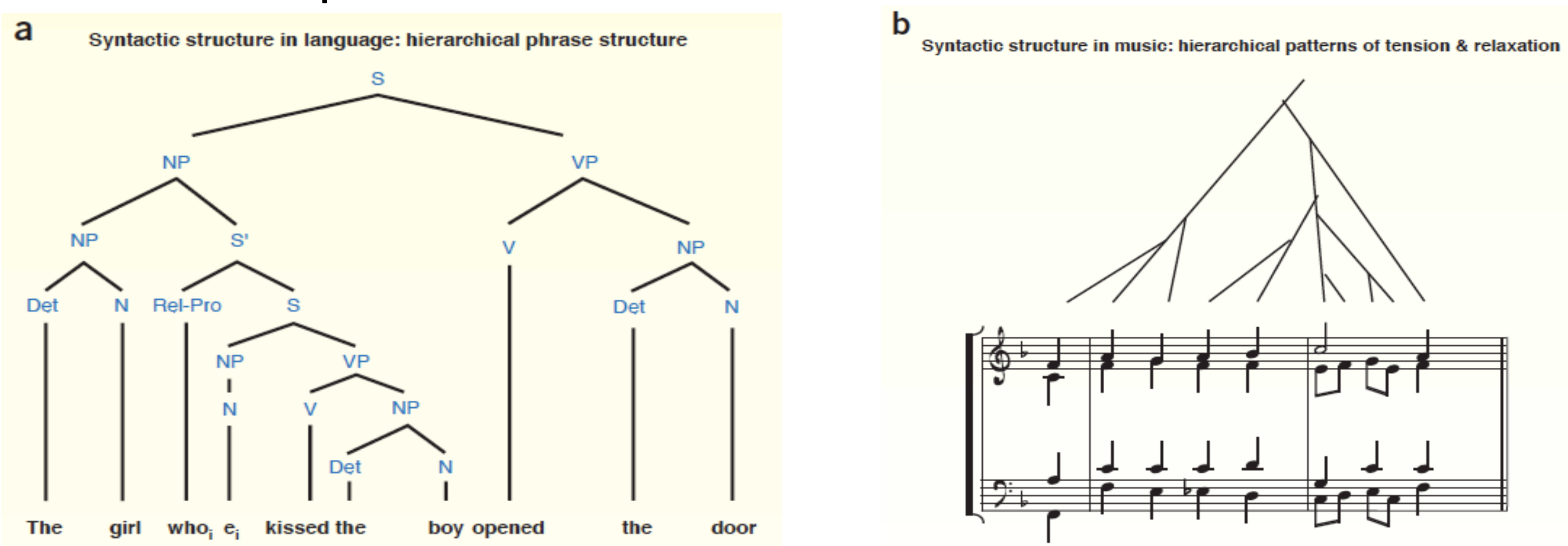
Figure 2. Between-subjects surface-based average showing greater response for song versus speech stimuli.

Cool, but ambiguous: does this reflect pitch or melodic contour?
Let's get serious…

Does Music Recruit Neural Machinery for Language?

Many have noted the commonalities between music and language:

- Both are distinctively human, natively auditory and unfold over time
- Both have complex hierarchical structure



Lots of claims of overlap between language and music from neuroimaging

But these are based on group analyses, which can find overlap even if it is not present in any individual subjects....

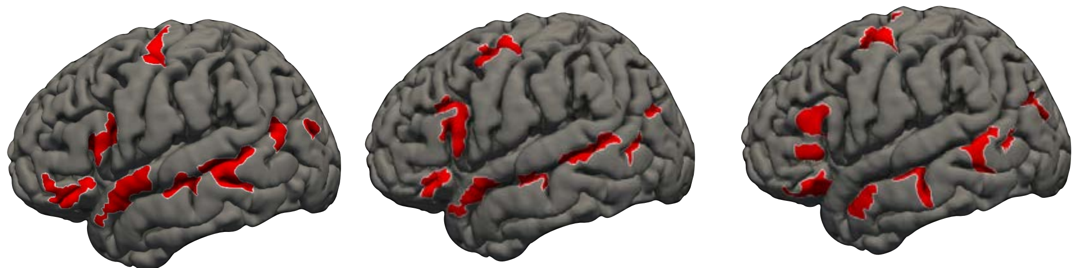
Luckily, however, Fedorenko did this right...

Does Music Recruit Neural Machinery for Language?

Ev Fedorenko

1. Functionally identify language regions in each subject individually

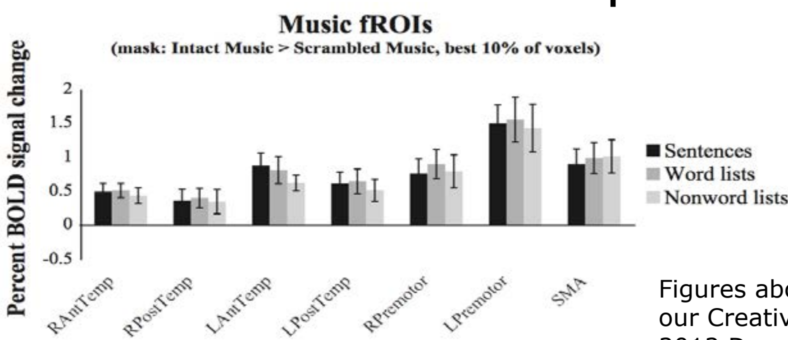
Sentences>
Nonwords
in 3 subjects



Then measure their response to intact and scrambled music.
No significant response in any language regions to intact>scrambled music.

2. Functionally identify candidate "music regions" (intact > scrambled music).

Then measure their response to language.



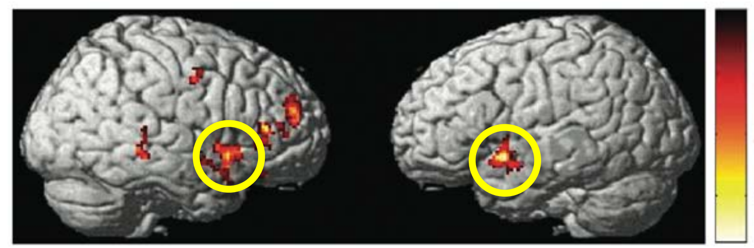


Fig. 6. Activation map from the random effects analysis for the Intact Music > Scrambled Music contrast (thresholded at $P < 0.001$, uncorrected) projected onto the single-subject template brain in SPM (single_subj_T1.img).

No higher response to sentences> nonwords.
Double dissociation of language and music.

Fedorenko, Behr, & Kanwisher, 2011

Fedorenko, et al., 2012

Does Music Recruit Neural Machinery for Language? *No!*

At least not the machinery for high-level language processing, that computes the meaning of a sentence independent of modality.
But what about mechanisms for speech perception?
Or, other auditory processing machinery?
Organization of auditory cortex not well understood.

What is the Functional Organization of Human Auditory Cortex?


Sam Norman-Haignere


Josh McDermott

Map of Frequency

Less consensus:
- **Speech regions**
- **Pitch Regions**
- **Spectrotemporal modulation?**
- **"Voice regions" ??**
- **"Music activations" ???**

Subsequent evidence:
Same regions also respond to pitch.

There is a problem with the scattershot approach the field has been taking:
ad hoc hypotheses
each tested with just a few stimuli
What if the main organization is
not something we would think to test?
We tried a new, data-driven approach….

Norman-Haignere, Kanwisher & McDermott, *Neuron*, 2015

Scan people while they listen to 165 of the most commonly-heard recognizable natural sounds (each is 2 seconds):

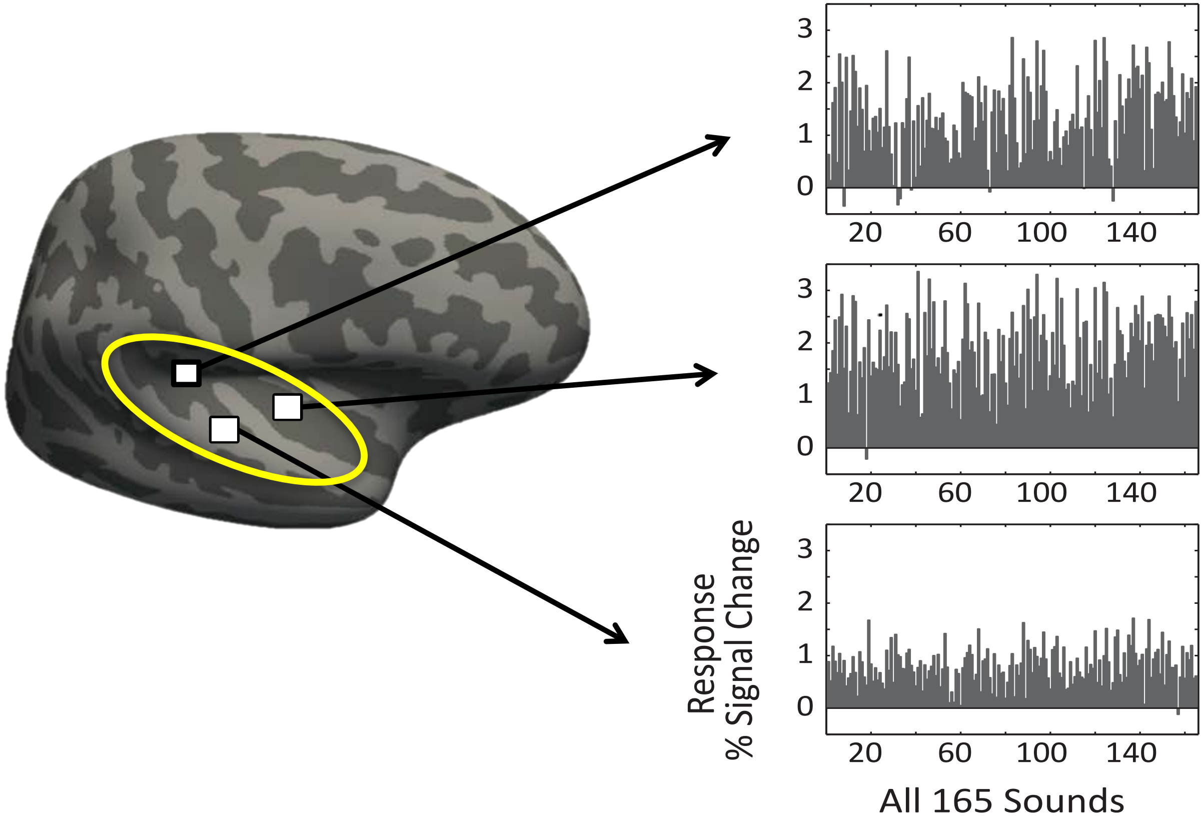
1. Man speaking
2. Flushing toilet
3. Pouring liquid
4. Tooth-brushing
5. Woman speaking
6. Car accelerating
7. Biting and chewing
8. Laughing
9. Typing
10. Car engine starting
11. Running water
12. Breathing
13. Keys jangling
14. Dishes clanking
15. Ringtone
16. Microwave
17. Dog barking
18. Walking (hard surface)
19. Road traffic
20. Zipper
21. Cellphone vibrating
22. Water dripping
23. Scratching
24. Car windows
25. Telephone ringing
26. Chopping food
27. Telephone dialing
28. Girl speaking
29. Car horn
30. Writing
31. Computer startup sound
32. Background speech
33. Songbird
34. Pouring water
35. Pop song
36. Water boiling
37. Guitar
38. Coughing
39. Crumpling paper
40. Siren
41. Splashing water
42. Computer speech
43. Alarm clock
44. Walking with heels
45. Vacuum
46. Wind
47. Boy speaking
48. Chair rolling
49. Rock song
50. Door knocking

Fairly comprehensive: Most sounds you would think of are on the list

Norman-Haignere, Kanwisher & McDermott, *Neuron*, 2015

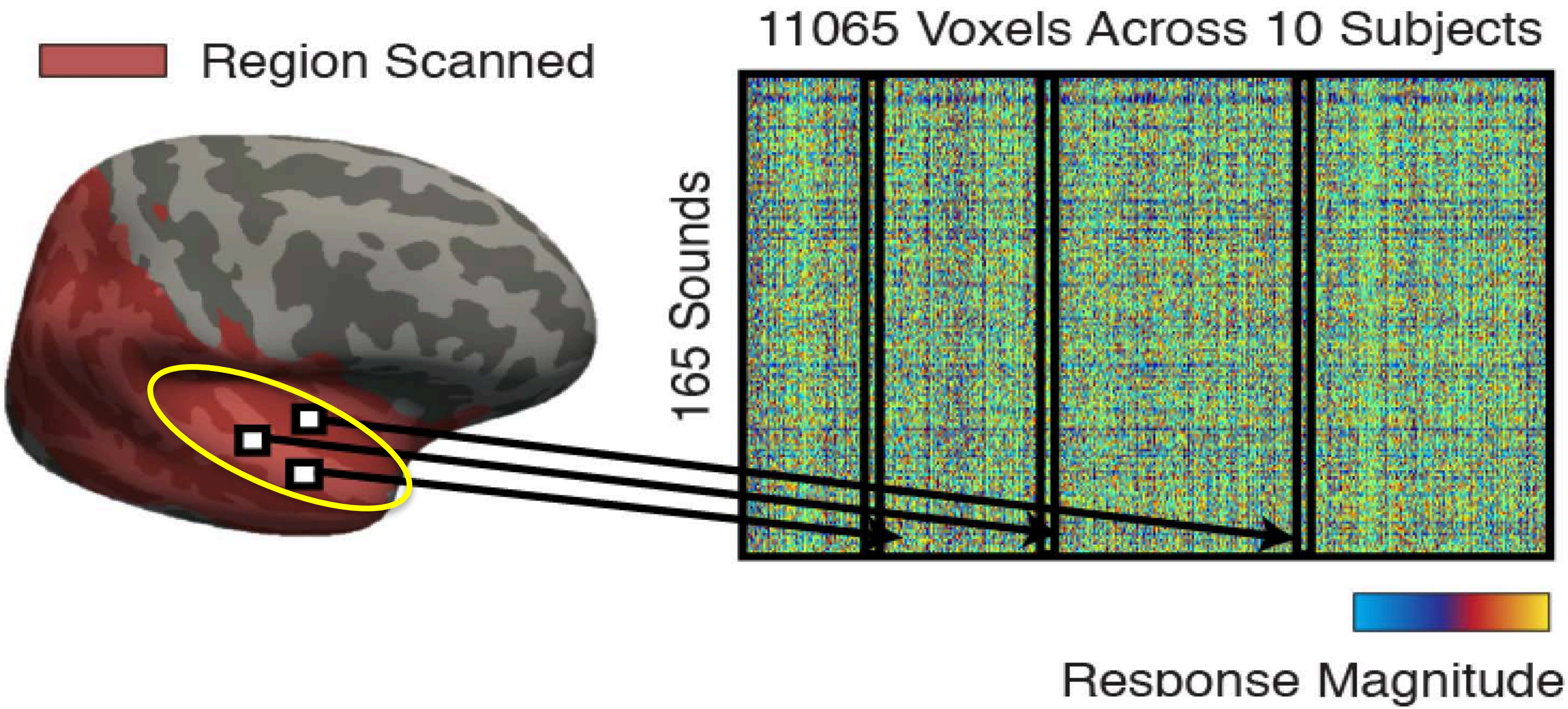
Voxel Responses

- For each voxel, we measure it's response magnitude to each sound



Do this for each voxel in auditory cortex for each of 10 subjects:11,065 voxels

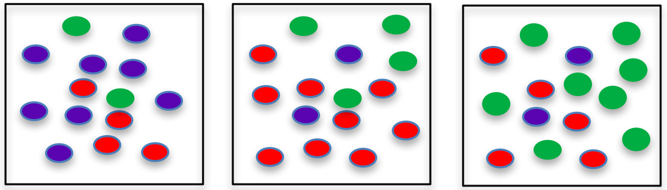
Data Matrix



Next: we do some math (~ICA) that tries to discover the basic structure in this array
Specifically…..

Modeling Assumptions

1. Voxel responses reflect the mixture of neural populations:

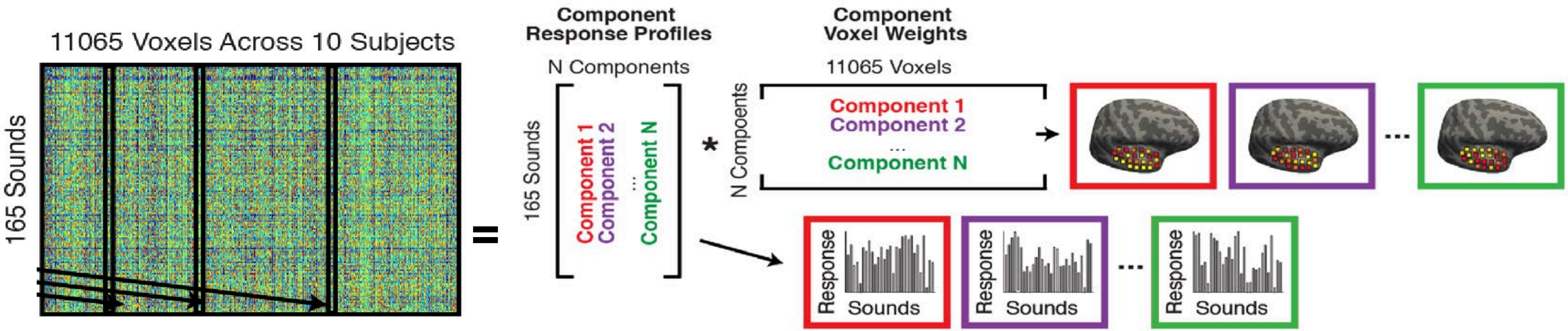


2. Each population has a canonical response profile across the 165 sounds

3. Voxel responses are the sum of the neural populations in each voxel

Goal: discover these canonical response profiles ("components")

Matrix Decomposition & ICA



Figures courtesy Elsevier, Inc., https://www.sciencedirect.com. Used with permission. Source: S Norman-Haignere, N Kanwisher, J McDermott, *Neuron* Vol. 88 (6) 1281-1296, December 16, 2015. https://doi.org/10.1016/j.neuron.2015.11.035

Factor response matrix into set of N components, each with:

- Response profile across the165 sounds
- Voxel weights specifying the contribution of each component to each voxel

Use ICA to search for components w/ independent voxel weights

- No information about sounds or anatomy used in decomposition
- Hypothesis space is huge and unconstrained (> 2^{165})
- This method should discover *the main dimensions that account for variance in the response across voxels in this stimulus set*

Six Components Account for Most of the Data

Four reflected ~expected acoustic properties.
One = low frq, one = high: tonotopy!

Two did not:

Really?

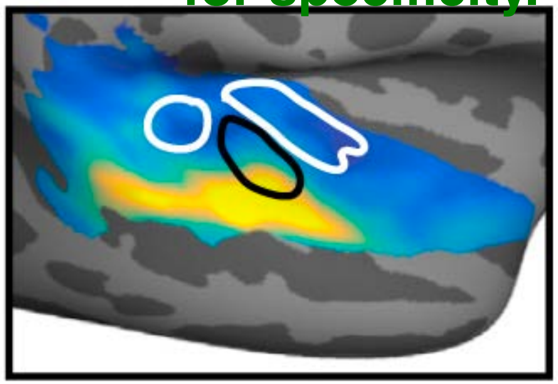
Speech

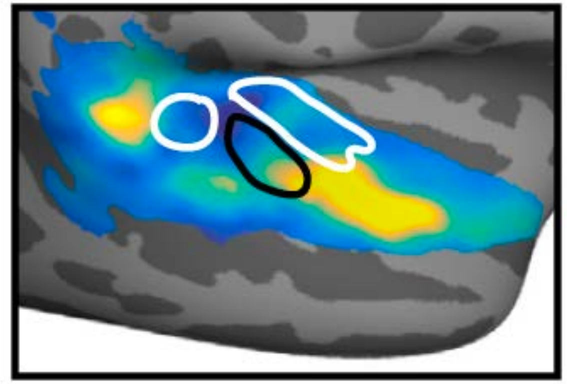
Component 5
Speech!

Not entirely new.
But strongest evidence for specificity.

Component : Music!

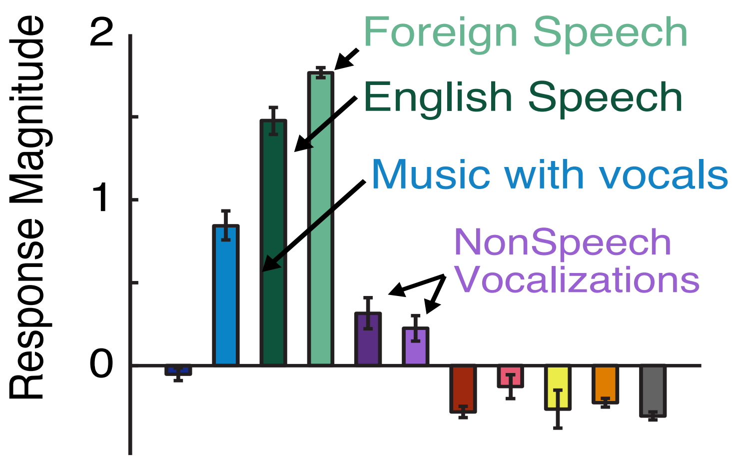
- **Double diss. of S & M**
- **Music does not just use mechs for speech**

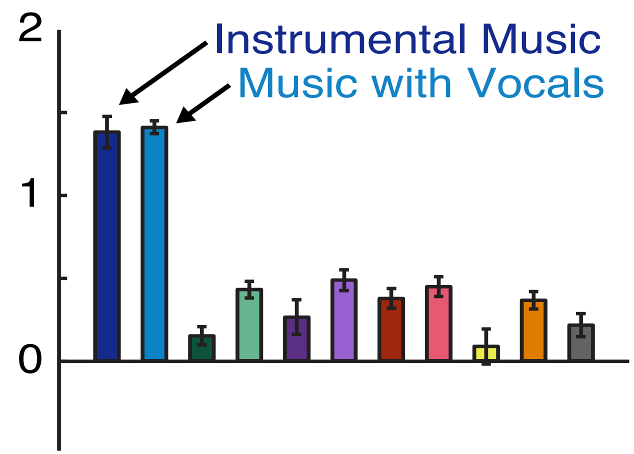




Sound Categories

- Instr. Music
- Vocal Music
- English Speech
- Foreign Speech
- NonSpeech Vocal
- Animal Vocal
- Human NonVocal
- Animal NonVocal
- Nature
- Mechanical
- Env. Sounds





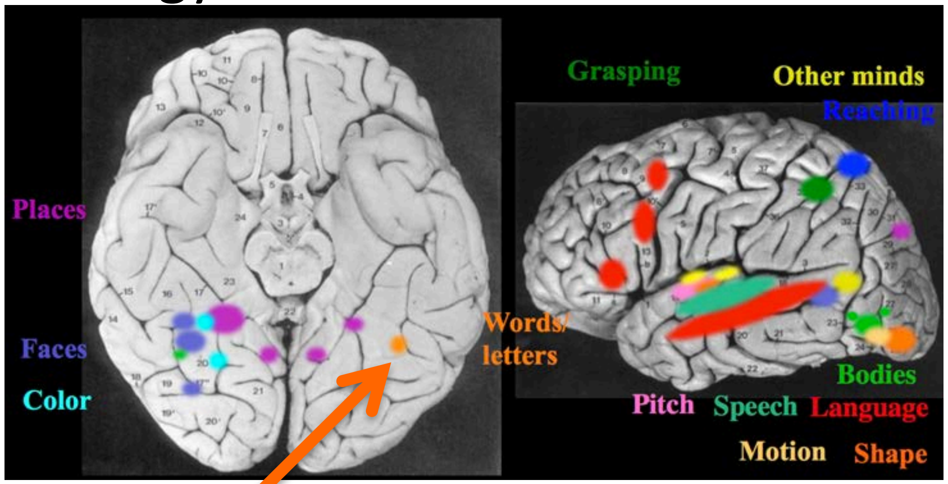
Music –Selective Component

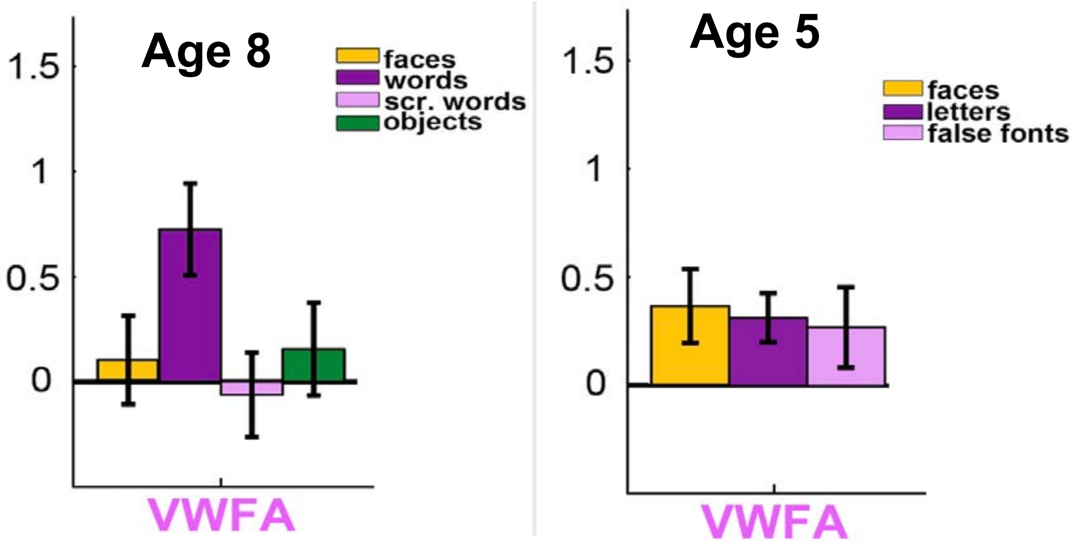
Really?

1. Can we replicate Sam's results?
   scan 20 new subjects
2. Is the music component a result of explicit training?
   10 people with ~ no explicit musical training
   10 highly trained musicians

Dana Boebinger
(& Sam & Josh)

Analogy: the Visual Word Form Area





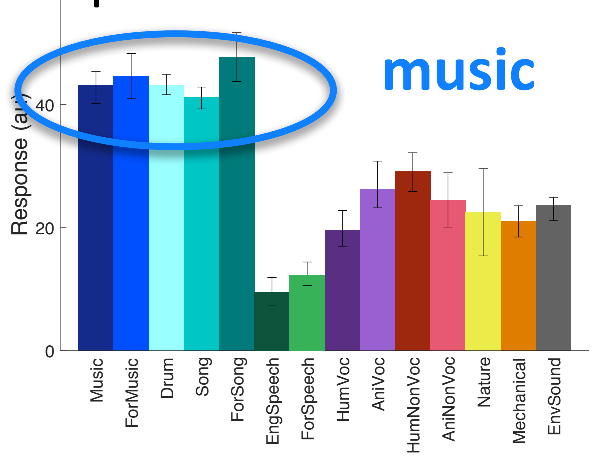
Is the music component like this?

Is Explicit Musical Training Required?

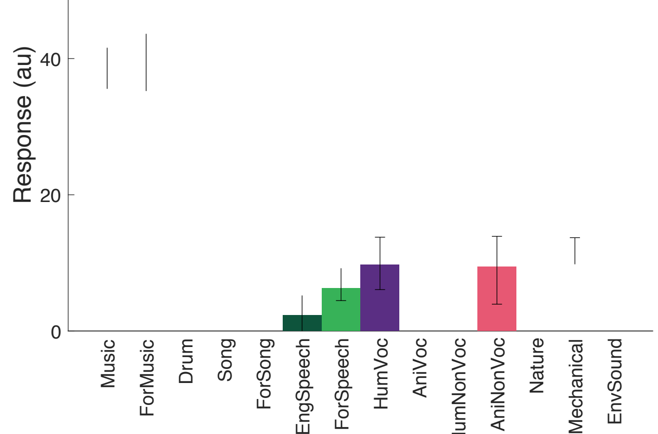
1. Can we replicate Sam's results? **Yes!**
   scan 20 new subjects
   same original stimuli, plus some new ones
2. Is music component evident even in ppl with no explicit training?

Dana Boebinger
(& Sam & Josh)

In 10 people with ~no explicit musical training



In 10 highly trained musicians



Music component is present in people with no explicit musical training.

Note this doesn't mean no *experience*, just no *explicit training*.

May be more selective/ have higher weights, in musicians than non (stay tuned).

Interrim Summary

1. Music perception does not engage cortical regions specialized for language understanding and vice versa.
2. Data-driven fMRI methods discover a strikingly music-specific component in human auditory cortex.
3. The music component does not respond to speech and vice versa.
4. It is present in people who have had no explicit musical training.
    so, not like the VWFA in requirement for explicit instruction
    (though, maybe like the VWFA in requirement for experience)

This is all very nice, but what is this “music component” anyway?
Presumably a population of neurons with this response profile.
But so far we have only inferred it mathematically.
We cannot directly observe this same selectivity in individual voxels.
Wouldn’t it be nice if we could observe it directly?
Perhaps, with a higher resolution method we could...

Intracranial Recordings

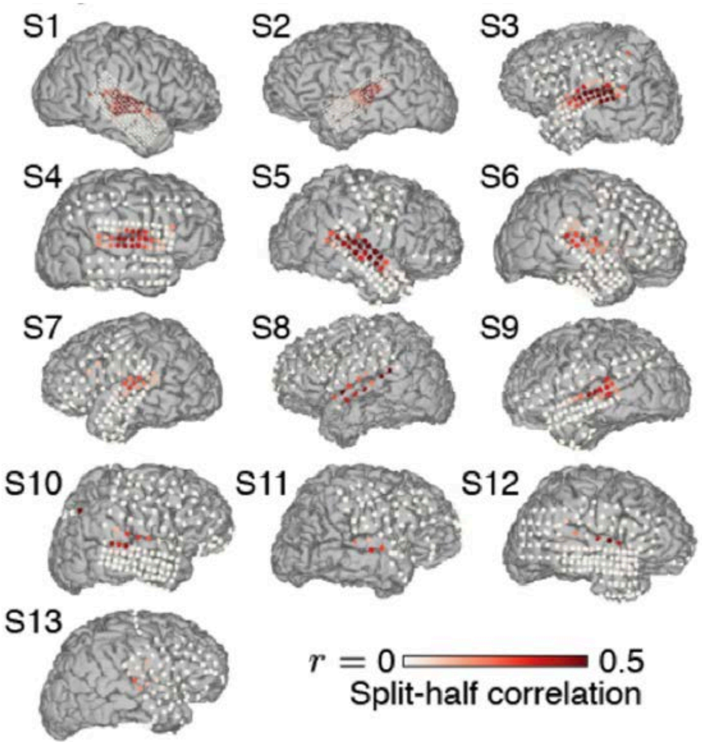
Sam Norman-Haignere

Gerwin Schalk

13 Neurosurgery patients w/ electrode grids

over superior temporal gyrus

271 electrodes w/ reliable responses



Measure high gamma responses of each electrode to the 165 sounds

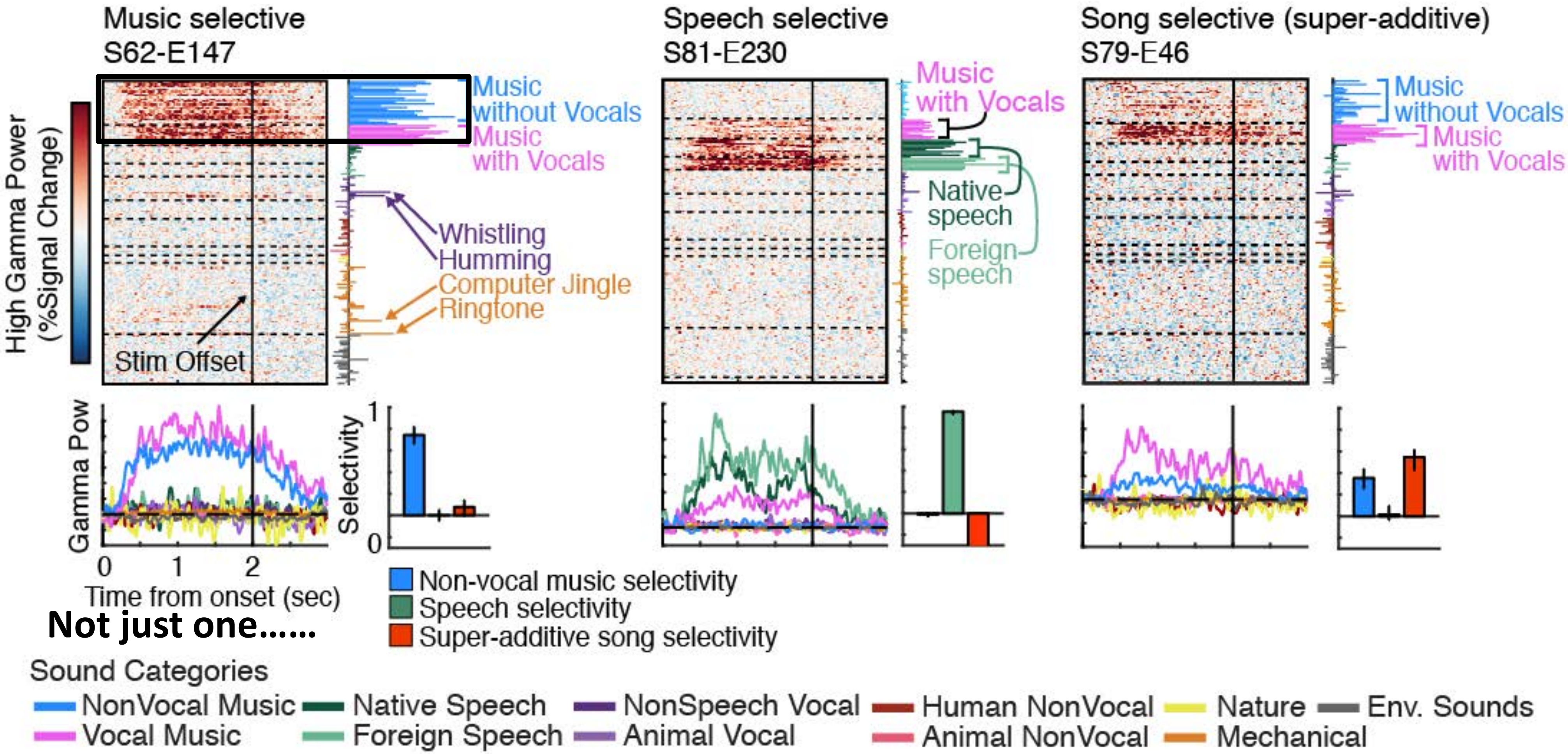
We find electrodes with three kinds of responses.

(Each electrode is categorized based on independent data.)

An example of the first kind of electrode…

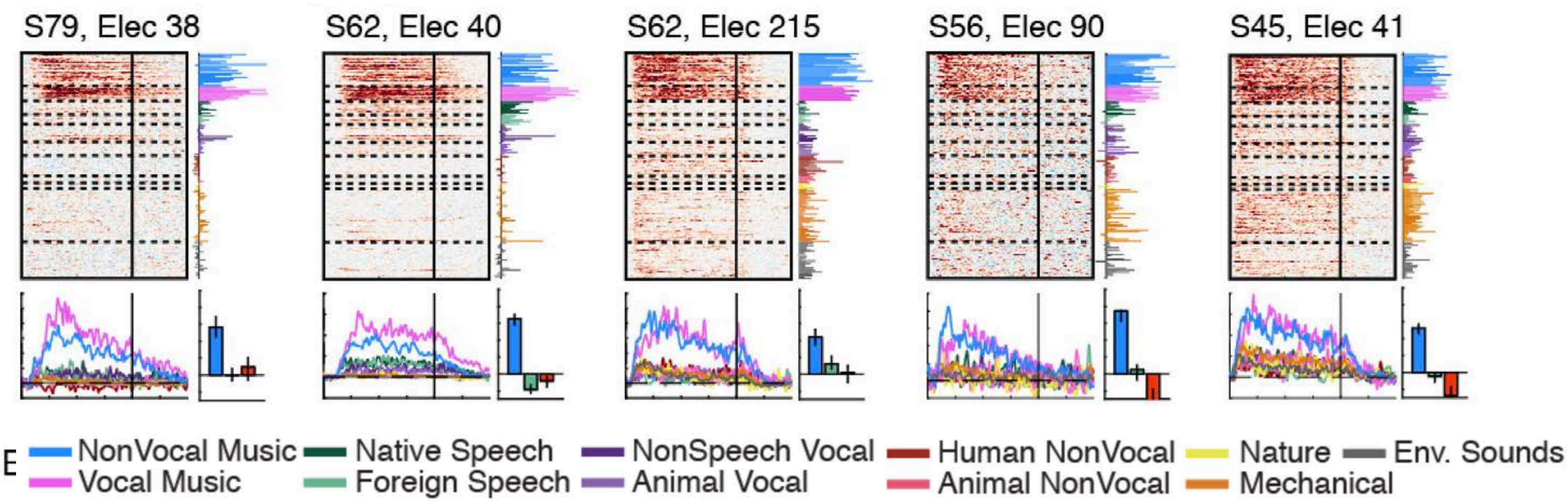
Intracranial Recordings

192 of these.
Nice.
But the real prize is….



Not just one……

Intracranial Recordings



Music Selectivity is for real!

Visible in raw gamma power in individual electrodes.

(Cannot see this in fMRI voxels.)

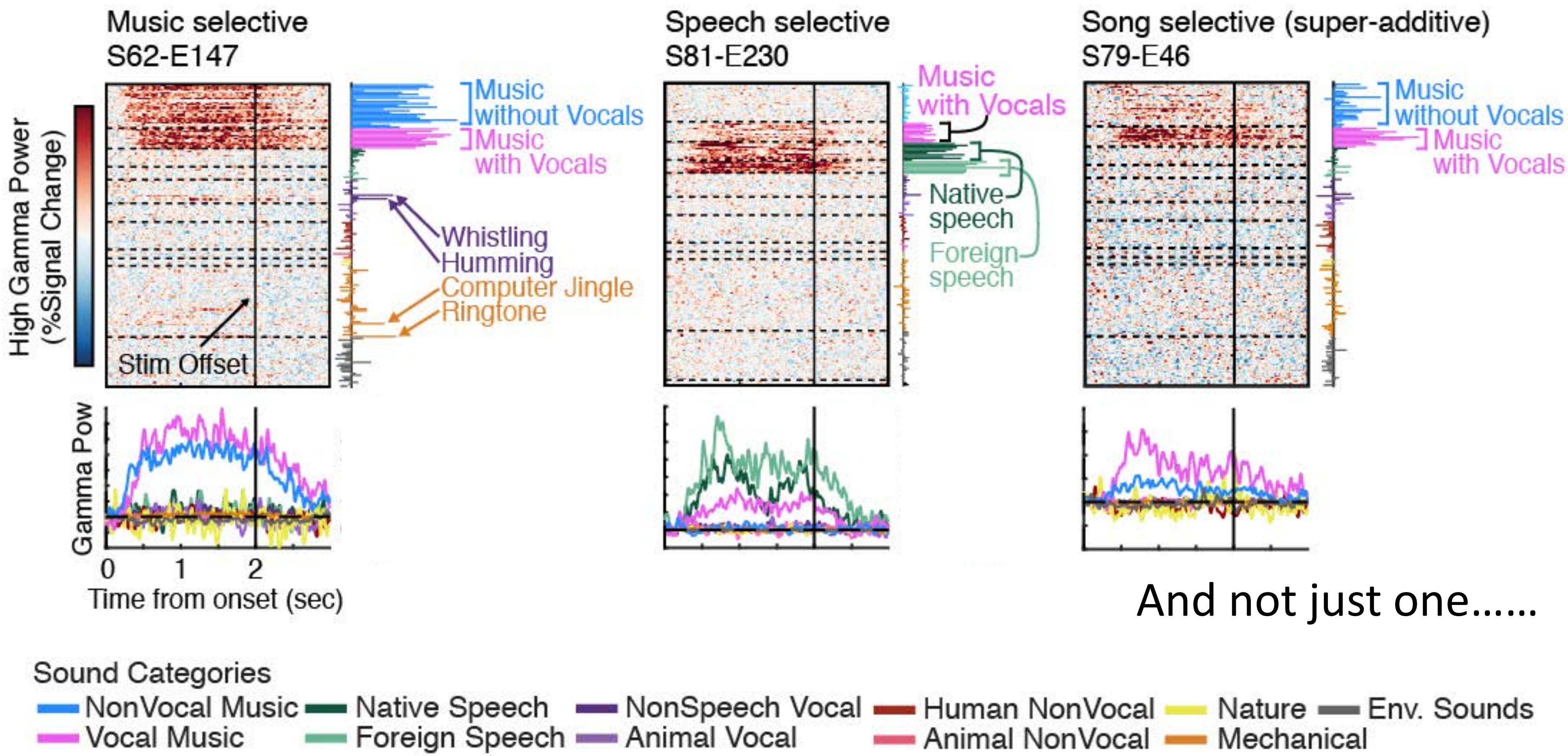
Validates the ICA method.

But we also found a surprise.

In addition to speech- and music-selective responses……

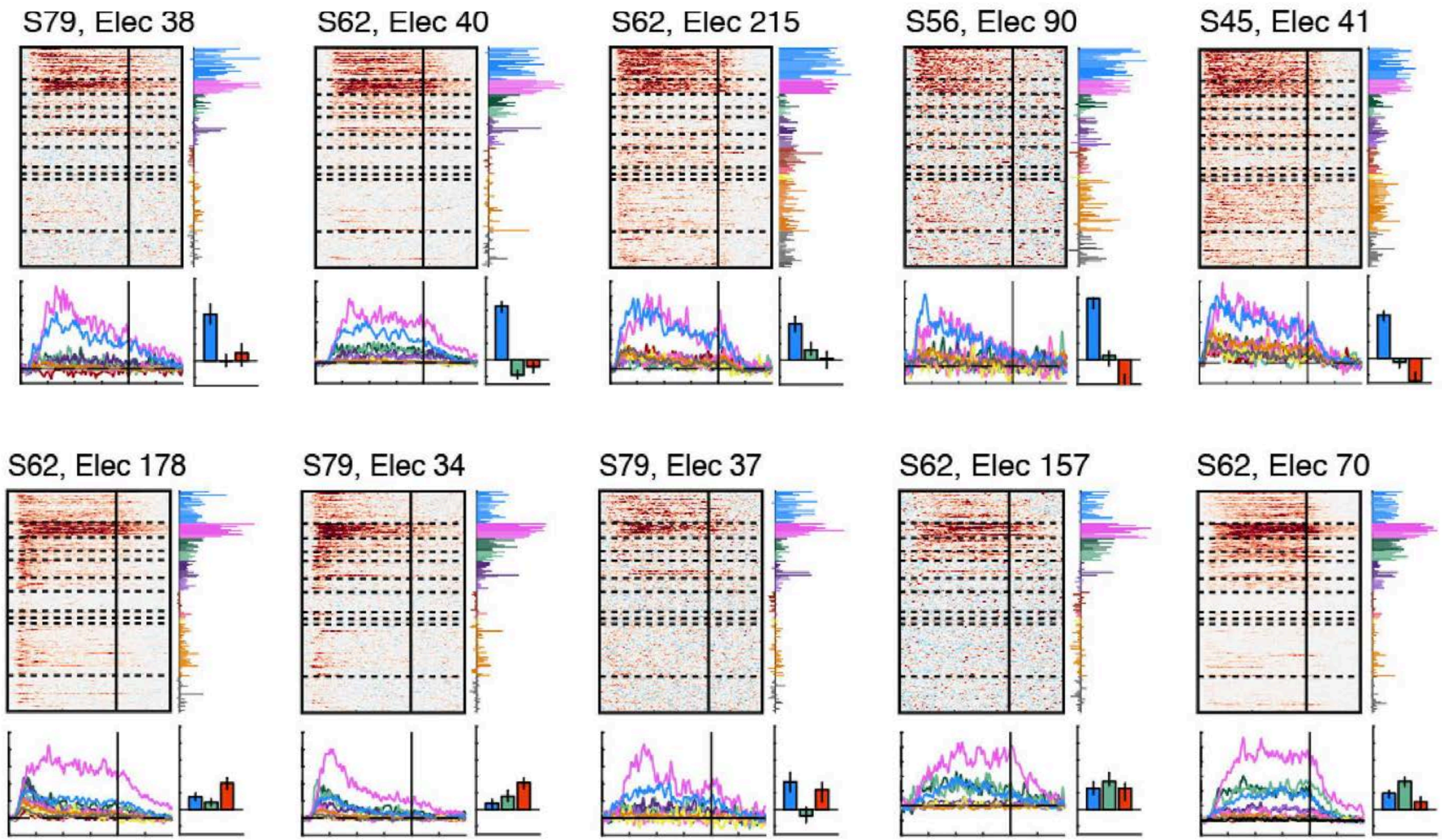
Intracranial Recordings

We also found……



And not just one……

Intracranial Recordings



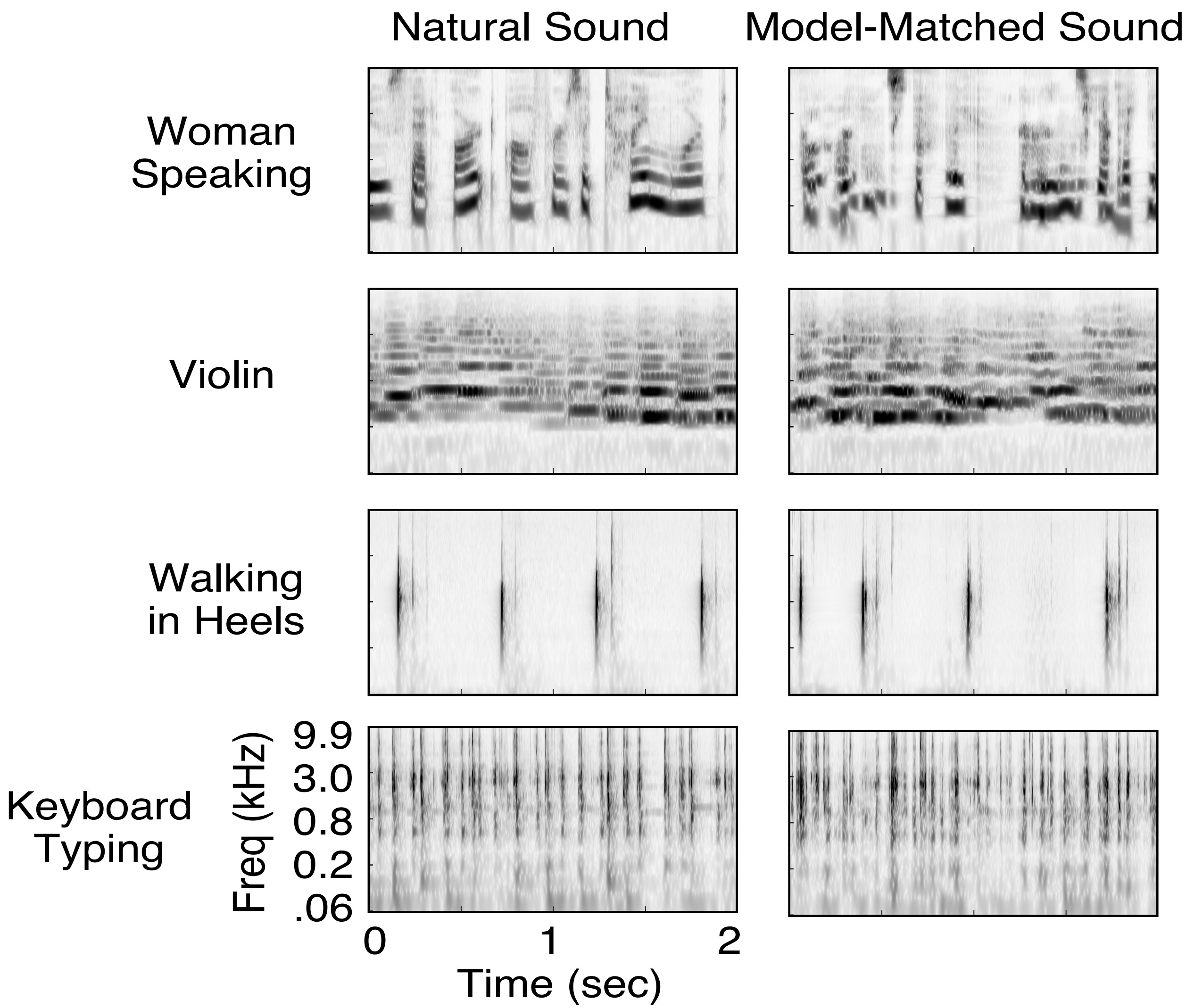
Song response is > sum of resp to music and speech (super additive).

Can these results be explained by low-level acoustic correlates?

Reminder: Norman-Haignere & McDermott (2018)

Idea: Create acoustically matched control stimuli using a standard model of A1 (linear spectrotemporal filters)

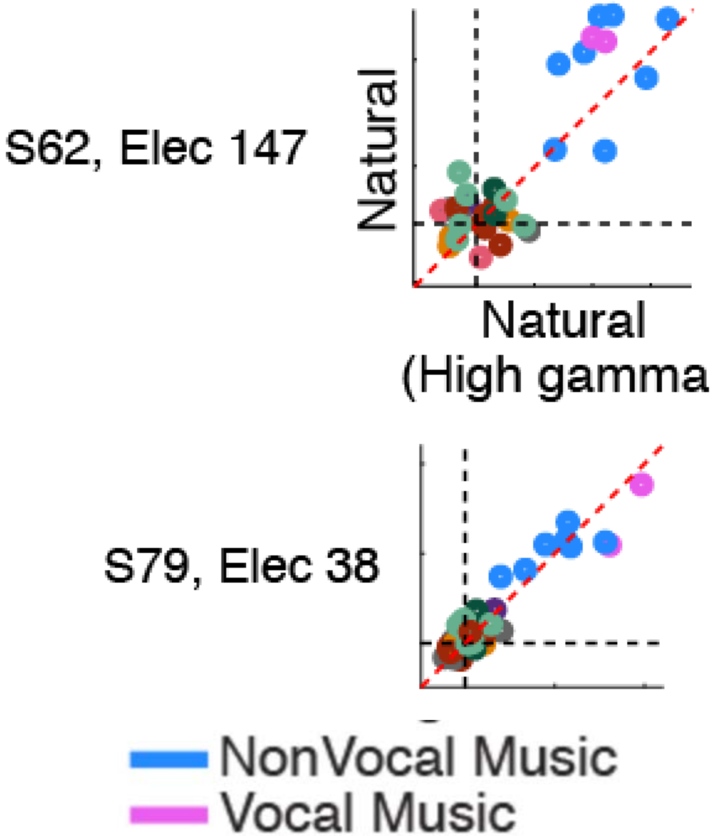
Indeed, A1 responds v similarly to original and model-matched version of stimulus.

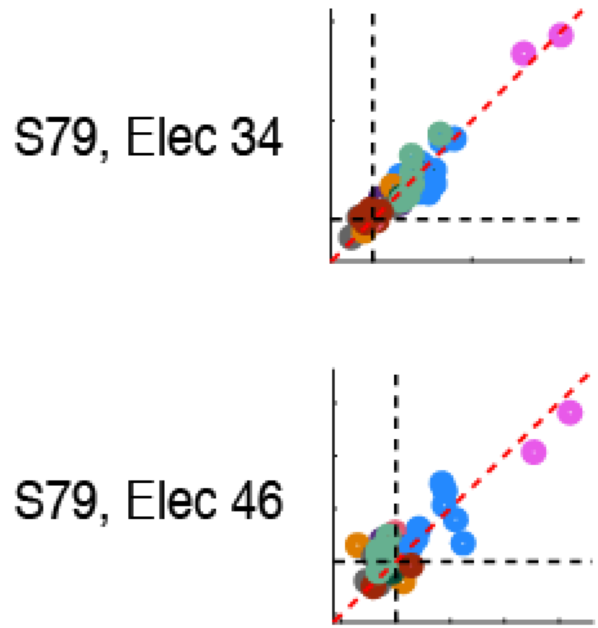


Intracranial Recordings

Two music-selective electrodes
showing test-retest reliability
Lower resp to "synthetic music"

Two song-selective electrodes
Lower resp to "synthetic song"





Can these results be explained by low-level acoustic correlates? No!

Conclusions

1. Music perception does not engage cortical regions specialized for language understanding and vice versa.
2. Data-driven fMRI methods discover a strikingly music-specific component in human auditory cortex.
3. The music component does not respond to speech and vice versa.
4. It is present in people who have had no explicit musical training.
   so, not like the VWFA in requirement for explicit instruction
   (though, maybe like the VWFA in requirement for experience)
5. The music-selective component inferred from fMRI now validated by direct recording from the surface of the brain.
6. New neural selectivity for vocal music discovered.
7. These selectivities cannot be accounted for by acoustic properties.

Raises so many questions...

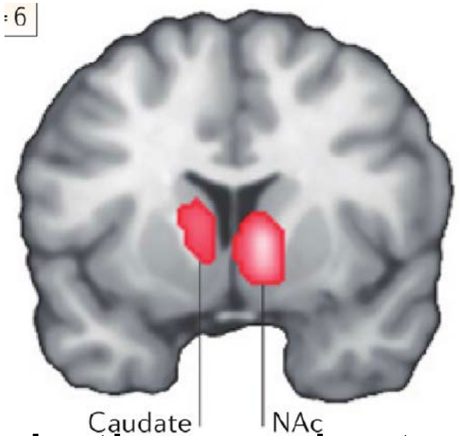
Open Questions

1. What features of music drive the music-selective responses?
   note-level structure (e.g. pitch and timbre)
   or patterning of notes (e.g. melodies, harmonies & rhythms)

2. How is music actually coded at the level of neural populations?

3. How is music and song selectivity constructed over development?
   Heather is working on this

4. Why is music *enjoyable*?



correlates of musical enjoyment = usual reward machinery: caudate and accumbens

but that doesn't tell us why

5. What if any part of this system is the product of natural selection?
   we still don't know!...

Major Puzzle: Why do we have Music?

Darwin's hypothesis: sexual selection: no evidence either way

"*[I]t 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*"

Mehr & Krasnow (2017): maybe

"*infant-directed song* [native form of music] *arose in an evolutionary arms race between parents and infants, stemming from the dynamics of parent-offspring conflict*" (infant needs to know parent is attending, parent has other needs)

Pinker (1994):

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"*

If so, it changes that machinery a lot over development, even with no explicit training

Put another way (common view):

~~Music is not an evolutionary adaptation at all,~~
~~but an alternate use of neural machinery that evolved to serve other functions,~~
~~like speech and language.~~