Male Talkers and **Female** Listeners

*What Can **Speech Responses** Tell Us about **Hearing**?*

Dr. Steve Aiken — Canadian Academy of Audiology Conference 2022

DALHOUSIE UNIVERSITY

Why is it always men talking and women listening?

Authors	Journal	Year	Stimulus	Female Listeners %
Easwar et al.	Eur J Neurosci	2018	**Male**	80
Bidelman	Neuroimage	2018	**Male**	71
Anderson et al.	J Acoust Soc Am	2013	**Male**	50
Krishnan et al.	Brain Lang	2011	**Male**	67
Aiken & Picton	Audiol Neurotol	2006	**Male**	50
Song et al.	J Cog Neurosc	2011	**Male**	76

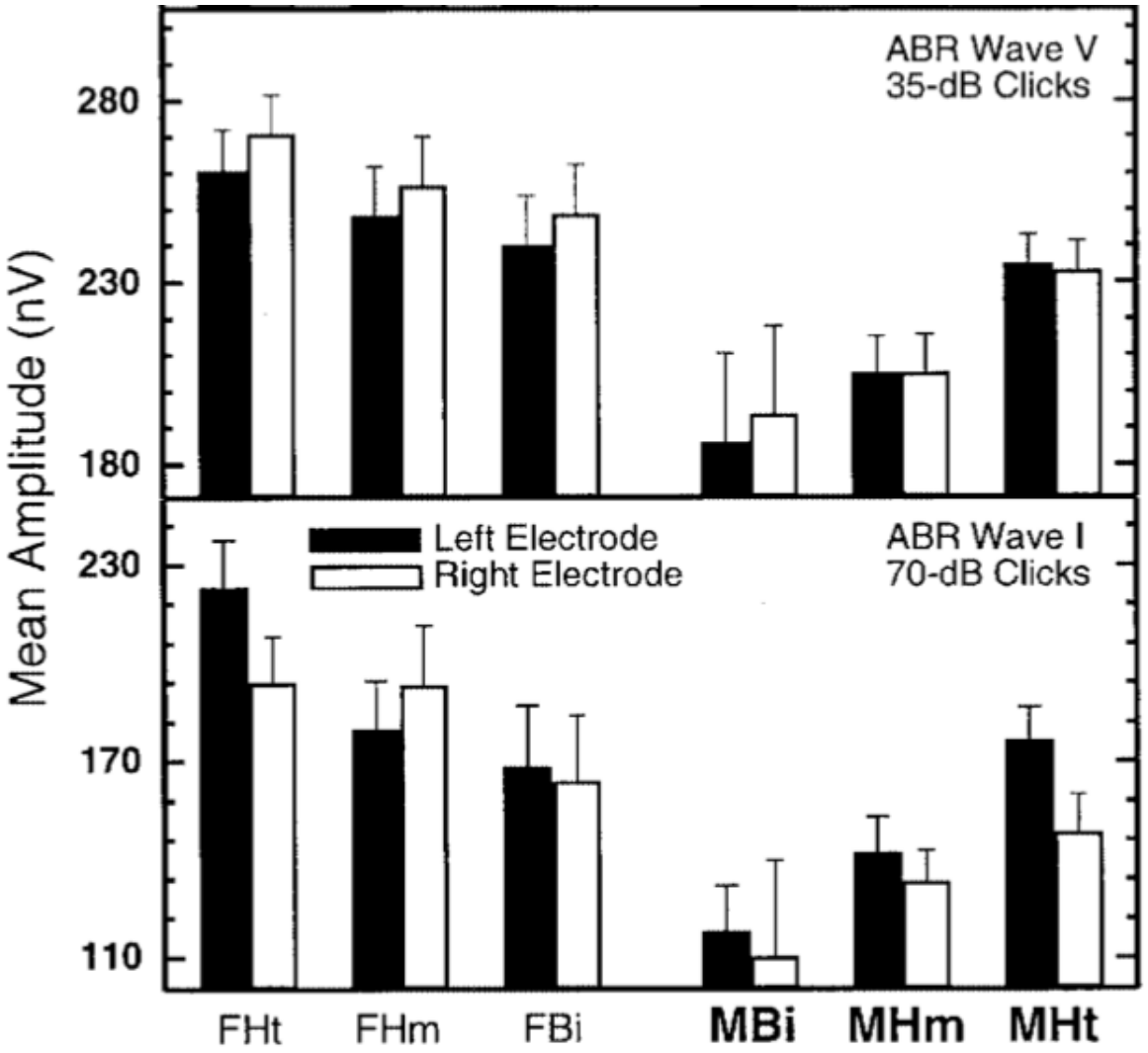
% Male Speakers	% Female Listeners
100	66

- more female listeners (66%)
- almost all male speakers (100%)

Why is it always women listening?

- women tend to give larger auditory evoked responses
  - e.g., McFadden & Champlin, 2000
  - females > responses than males
  - also relates to sexual orientation

- most people in audiology and neuroscience neuroscience labs are women!



Why is it always men talking and women listening?

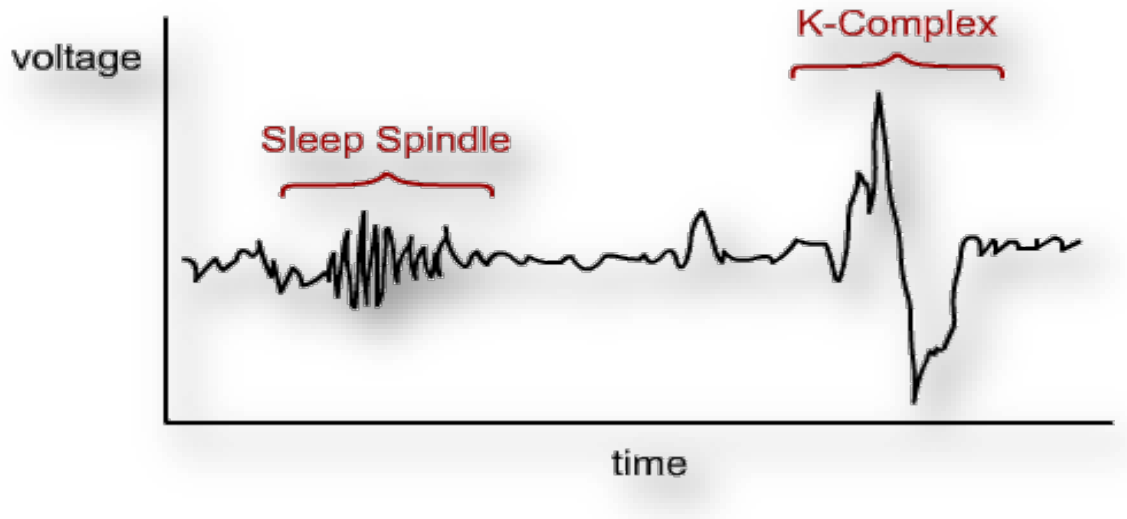
Authors	Journal	Year	Stimulus	Female Listeners %
Easwar et al.	Eur J Neurosci	2018	**Male**	80
Bidelman	Neuroimage	2018	**Male**	71
Anderson et al.	J Acoust Soc Am	2013	**Male**	50
Krishnan et al.	Brain Lang	2011	**Male**	67
Aiken & Picton	Audiol Neurotol	2006	**Male**	50
Song et al.	J Cog Neurosc	2011	**Male**	76

% Male Speakers	% Female Listeners
100	66

- more female listeners (66%)
- almost all male speakers (100%)

The First Auditory Evoked Response

- When stimulated with a loud sound, it is sometimes possible to see a change in the EEG
  - Pauline Davis first noticed that the EEG changed a bit when there were loud sounds (1939)
  - called this the K-Complex (“Knock”?)



Hallowell Davis and Pauline Davis

- together they were the father and mother of auditory evoked responses

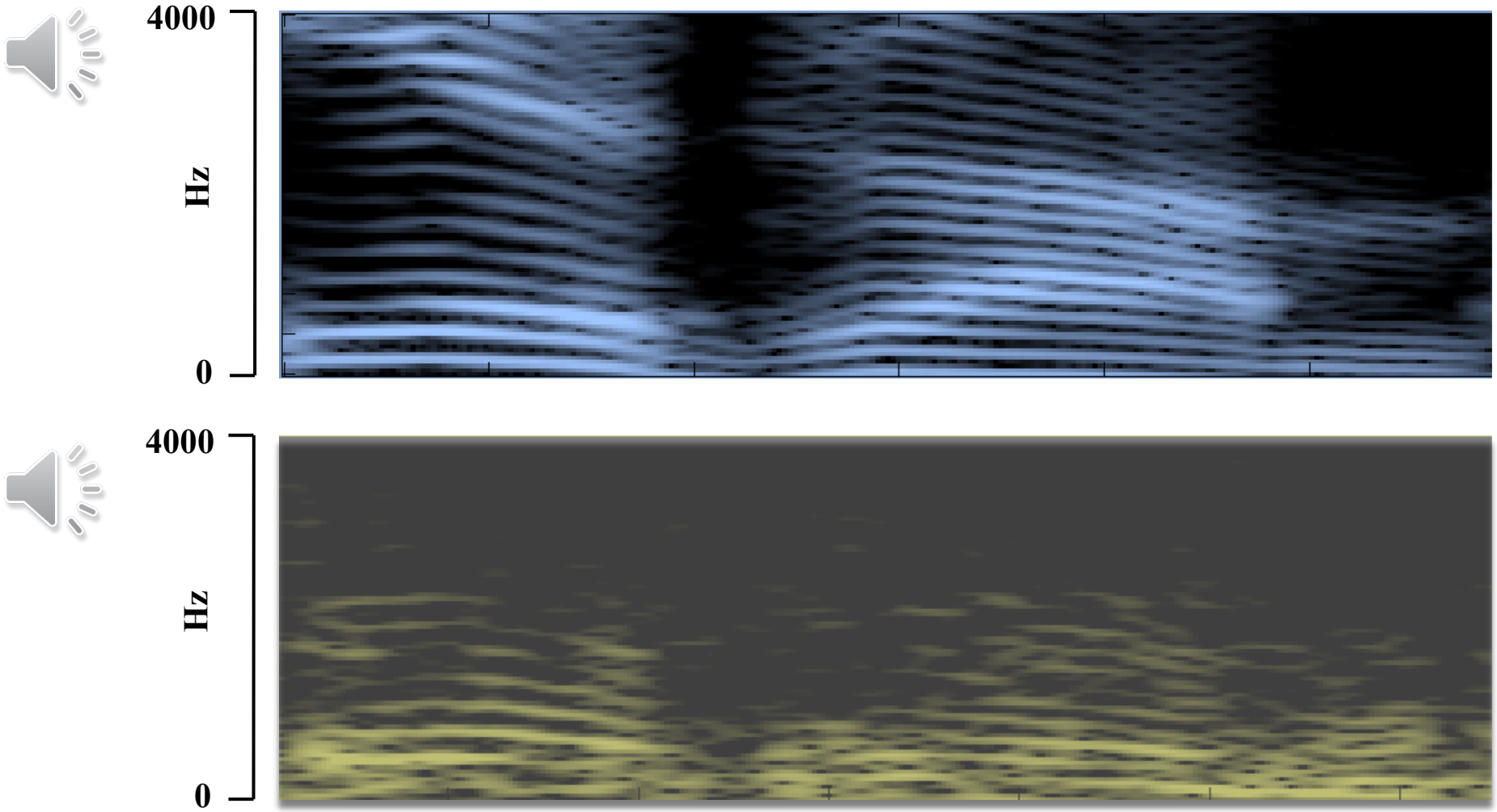
Pauline Davis

- **Davis, Pauline Allen** (Mrs. Hallowell), of Boston, Mass., signed with Near East Relief at Constantinople in March, 1921, and worked there until she was assigned to Derindje in April. During the Turco-Greek fighting at Ismid and Derindje she was slightly injured while bringing in wounded soldiers from between the firing lines. On September 18, 1921, she came back to America, but returned to Constantinople August 9, 1922. When the refugees poured into Constantinople and Greece after Smyrna Miss Allen was indefatigable both in the camps around Constantinople and in Athens, where she went in December, 1922. Back in Turkey she took over the department of general relief for 4,890 refugees in four camps. She contracted typhus in March, 1923, and was seriously ill for several weeks. When she was merely convalescent in April she helped in the rescue of 600 refugees who were too terrified to leave a burning building. Miss Allen led the bluejackets from the U.S. Destroyer "McCormick" through obscure parts of the building and every person was taken out safely. Captain Mannix of the Destroyer mentioned Miss Allen's pluck in a report to the Navy Department. In July, 1923, Miss Allen married Dr. Hallowell Davis, whose skill had pulled her through the attack of typhus. They reached America August 20, 1923, and are now living at 44 Edgehill Road, Brookline, Mass. Mrs. Davis is doing physiological research. ***(June 1924)***

Question of the Day

Why do we always use male speakers to evoke speech responses?

What are speech evoked responses?



The Auditory Nerve Phase-Locks to Inner Hair Cells

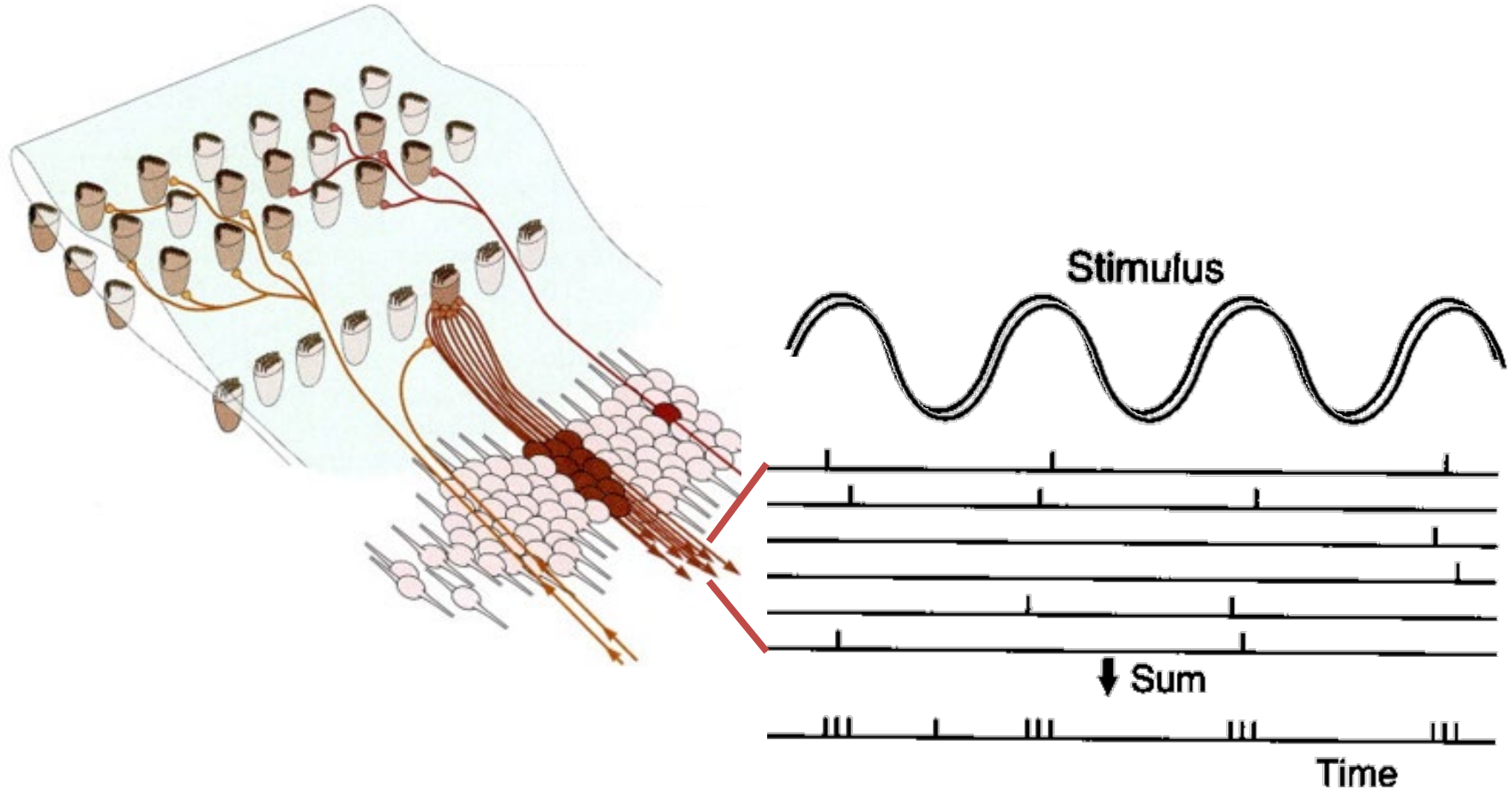
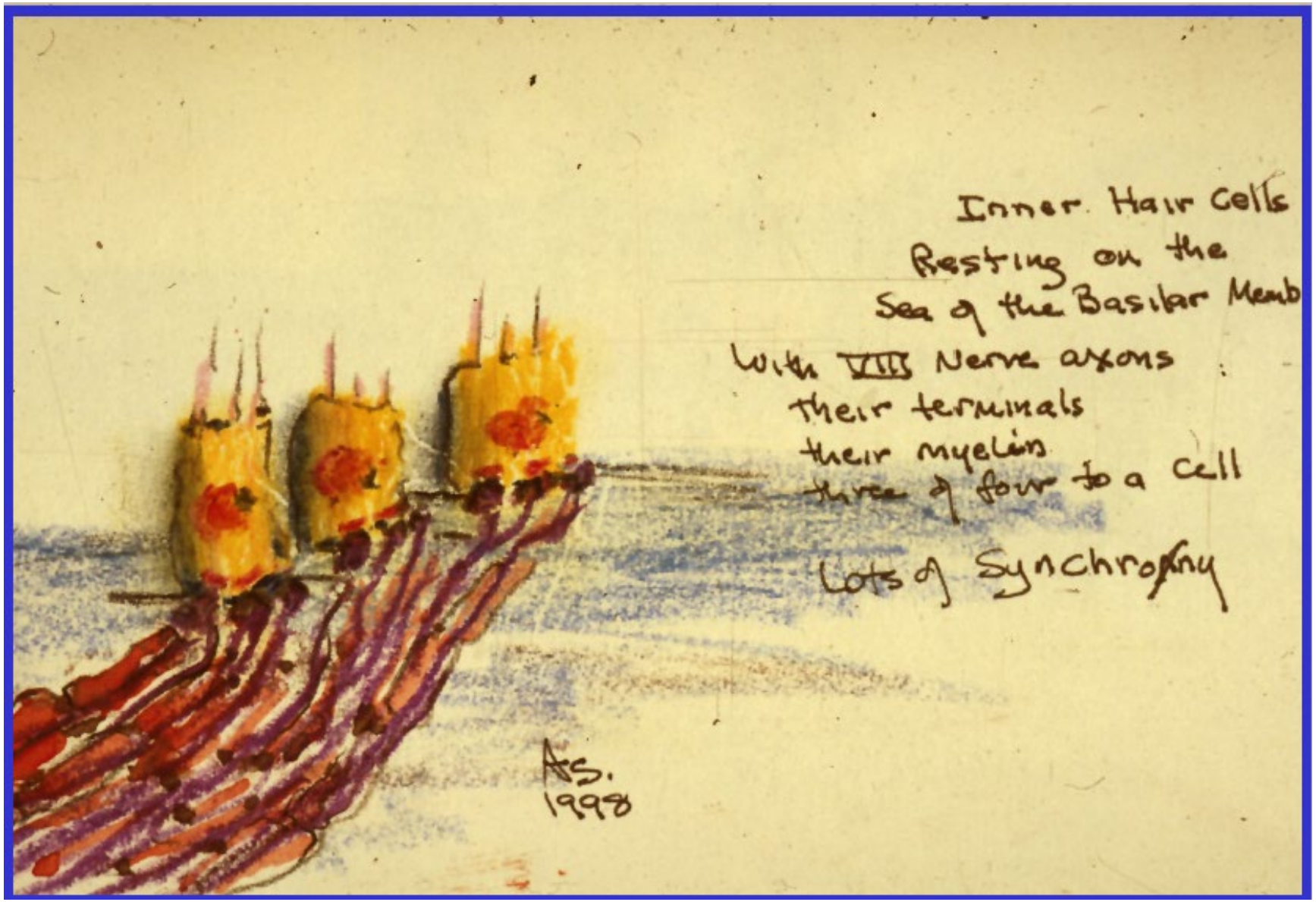
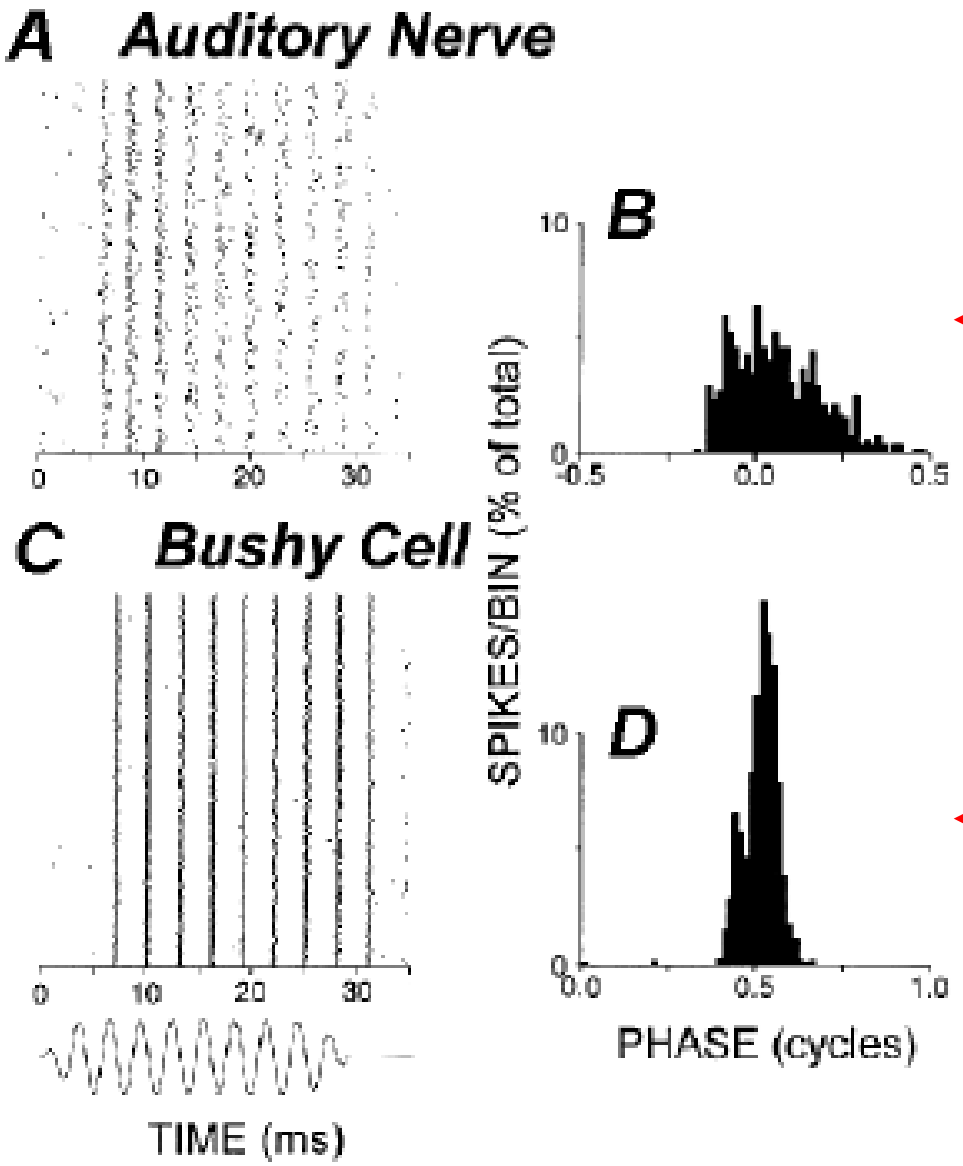


Figure source: Kandel ER. *Principles of Neural Science, Fourth Edition*. New York: McGraw Hill; 2000:602

Lots of Synchrony in the Auditory System



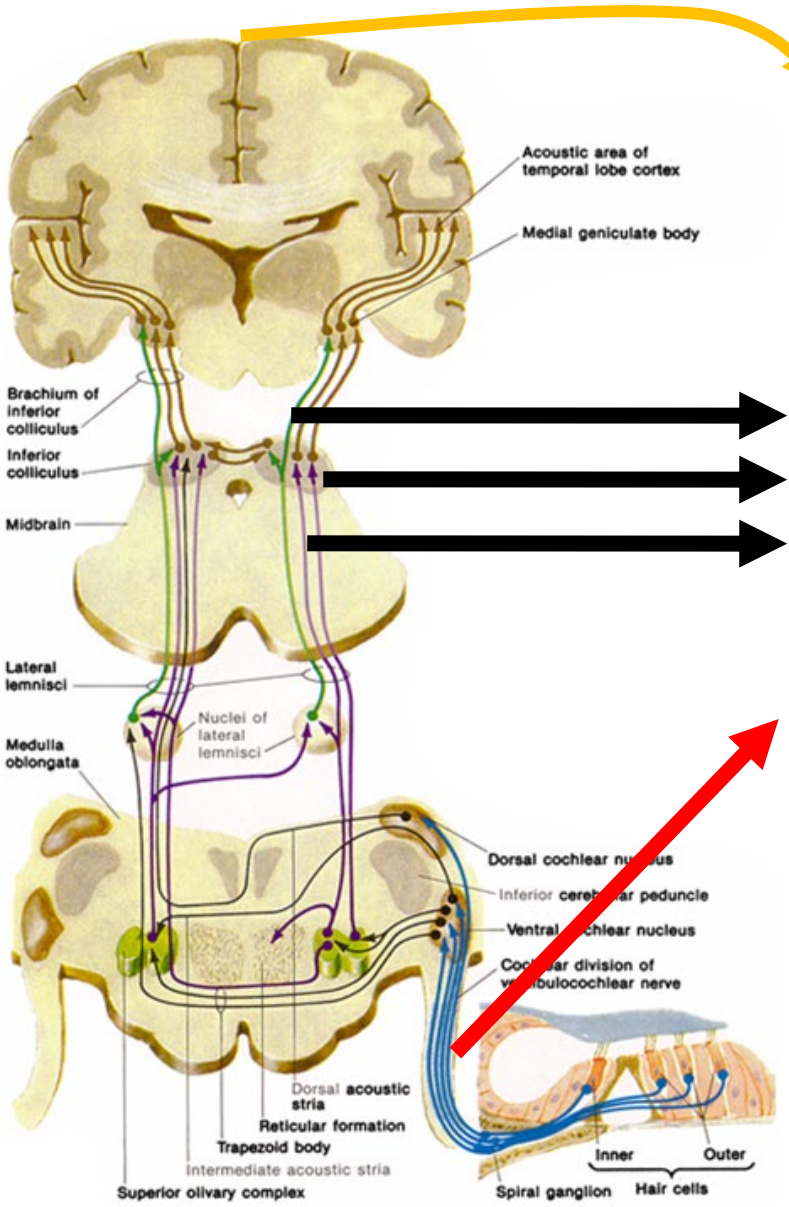
Phase-Locking in the AN and Cochlear Nucleus



- Phase-locking is **more precise** in the cochlear nucleus (e.g., bushy cells) than in the auditory nerve

Joris et al., 1994

The Frequency Following Response



Problems with the Frequency Following Response

- responses can only be recorded to ~1500 Hz
  - we need to know about hearing above 1500 Hz too!

- responses tend to be **VERY** small (e.g., 10 nanoVolts)
  - can be hard to record and buried in the noise

- responses are **VERY** easily contaminated with electrical artefact!
  - the stimulus and response have the same frequencies in them… so any artefact from the stimulus **will look just like the response**

Solution: ‘Envelope’ Following Responses!

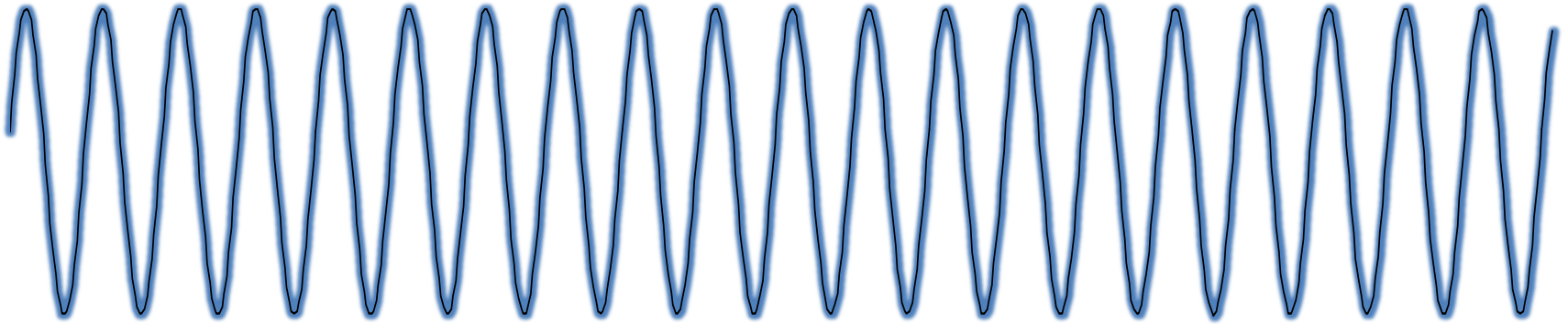
- let’s say I want to know how someone is hearing at 2000 Hz
  - this is too high in frequency to get a frequency following response
  - and even if I could, it would be tiny and easily confused with artefact

- what I need is:
  - a tone at 2000 Hz that is **amplitude-modulated** (turned up and down quickly) at some slower rate, where I can record larger responses from the brainstem!

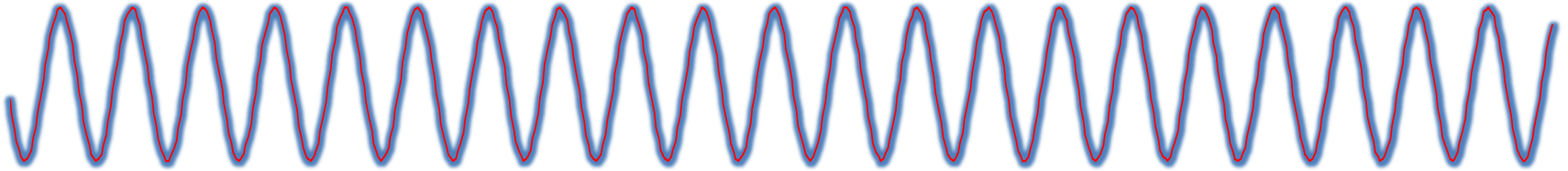
- these are also called auditory steady-state responses (ASSR)

Three Tones *Around* 2000 Hz (+/– 100 Hz)

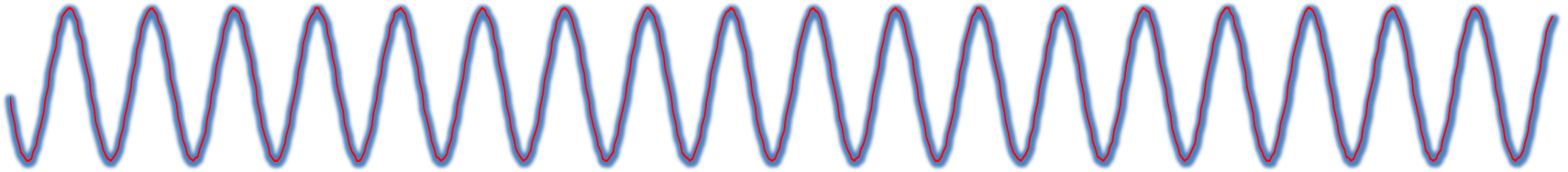
2000 Hz



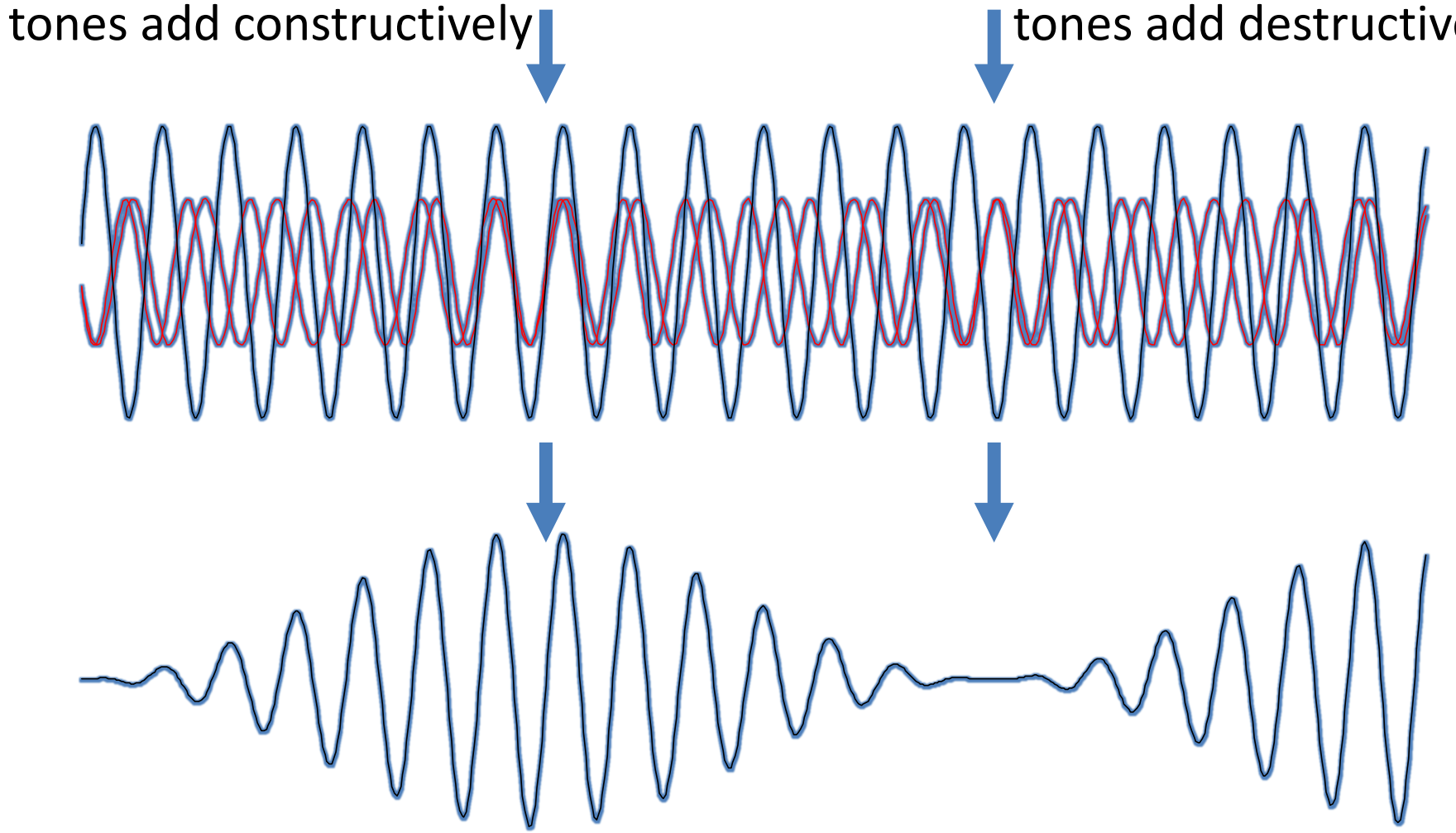
1900 Hz



2100 Hz

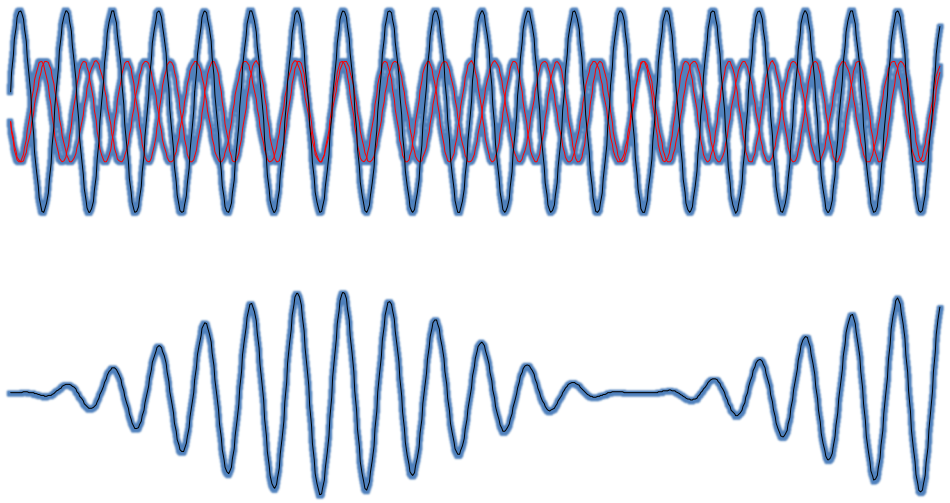


When these are added together…



The sum of these three tones around 2000 Hz will be a 2000 Hz tone that is amplitude modulated at 100 Hz

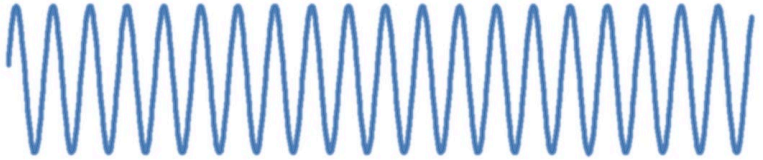
Envelope-Following Responses (or 'Auditory Steady State Responses)



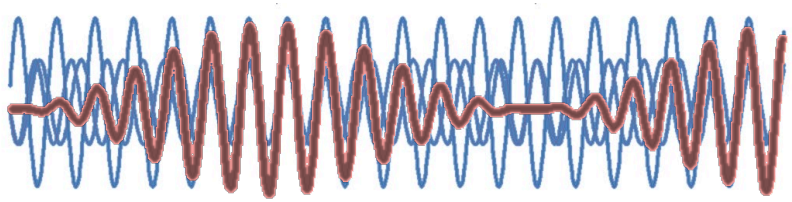
- these are responses to <u>fluctuations in amplitude</u>
- e.g.,
  - tones go in around 2000 Hz (and thus test hearing around 2000 Hz)
  - the brainstem locks to the fluctuation that you get when you add the tones together (e.g., 100 Hz)

EFR/ASSR Example with 1000 Hz, Modulated at 82 Hz

1000 Hz tone



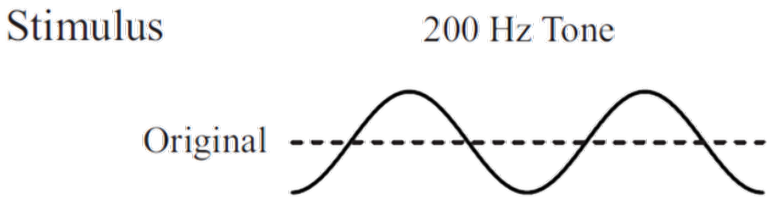
1000 Hz +/– 82 Hz

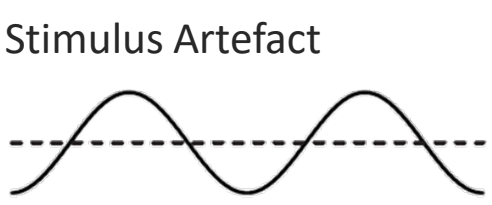


- Why is this better?
  - no problem with the 1500 Hz limit in the brainstem
  - responses are larger at low frequencies — easier to separate from noise
  - easy to separate from artefact!

Separating EFR/ASSR from Artefacts: Alternating Polarity

- traditional solution is to present the stimulus in two polarities (up/down), and then average responses
- stimulus artefact goes away, EFR/ASSR does not go away!

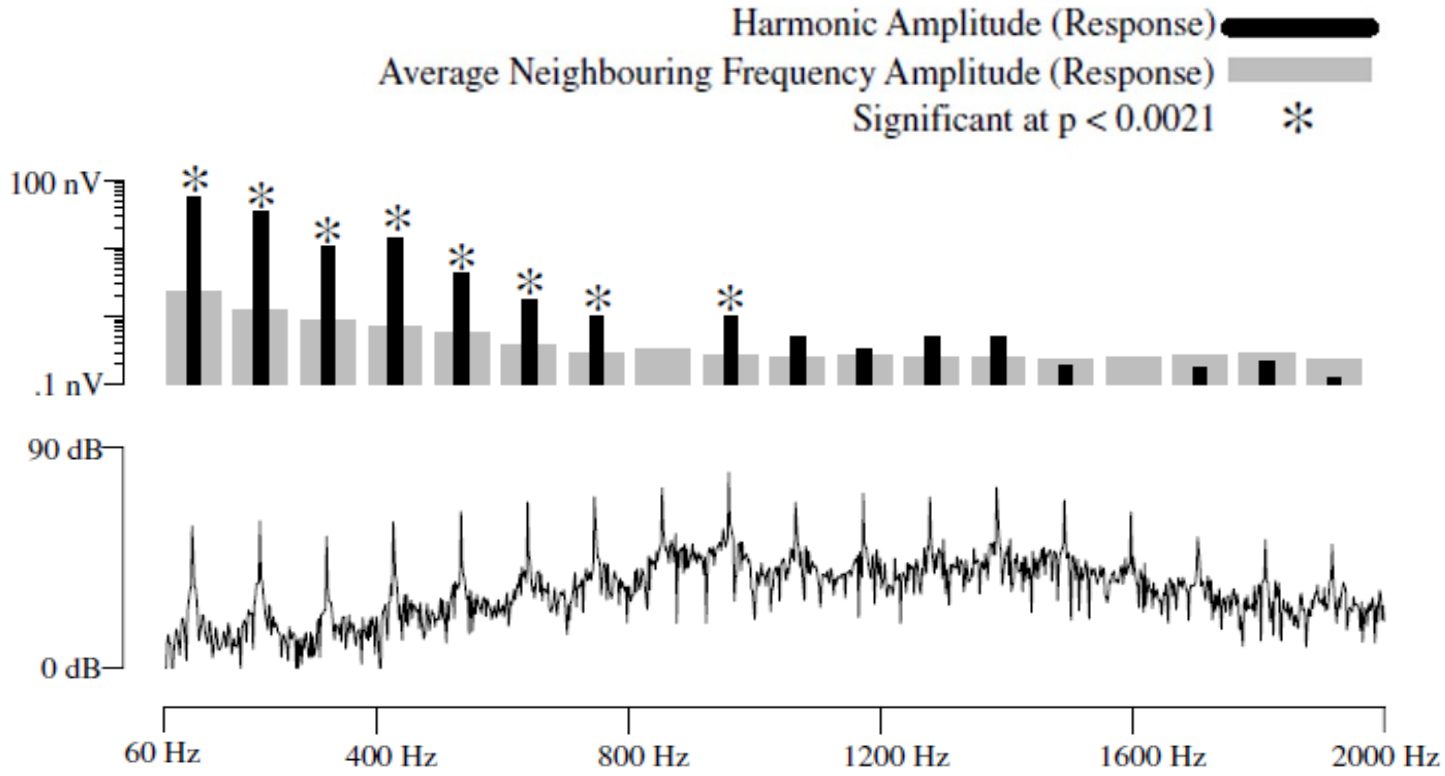




EFR/ASSR to a Speech Sound (/a/ vowel)

EFR/ASSR

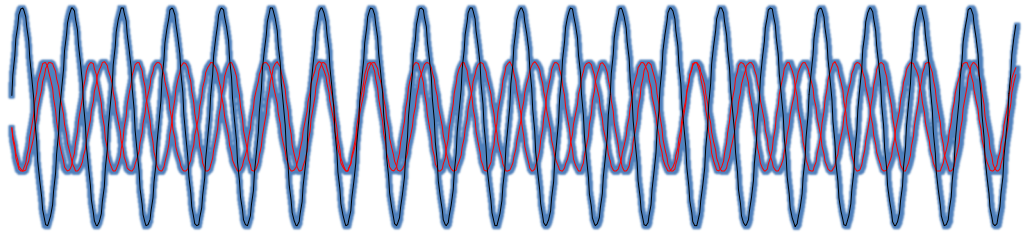
an /a/ vowel

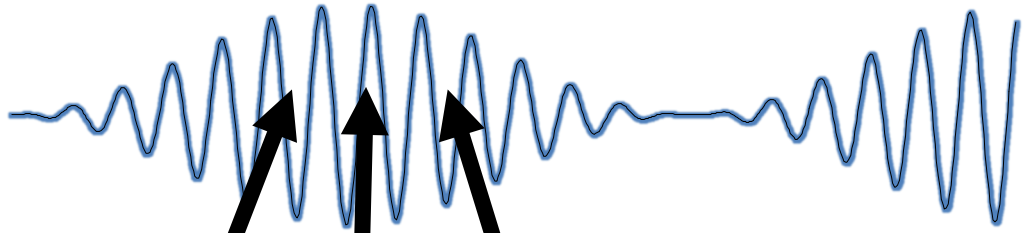


Aiken & Picton, Hear Res, 2008

Why such large responses at the lowest (fundamental) frequency?

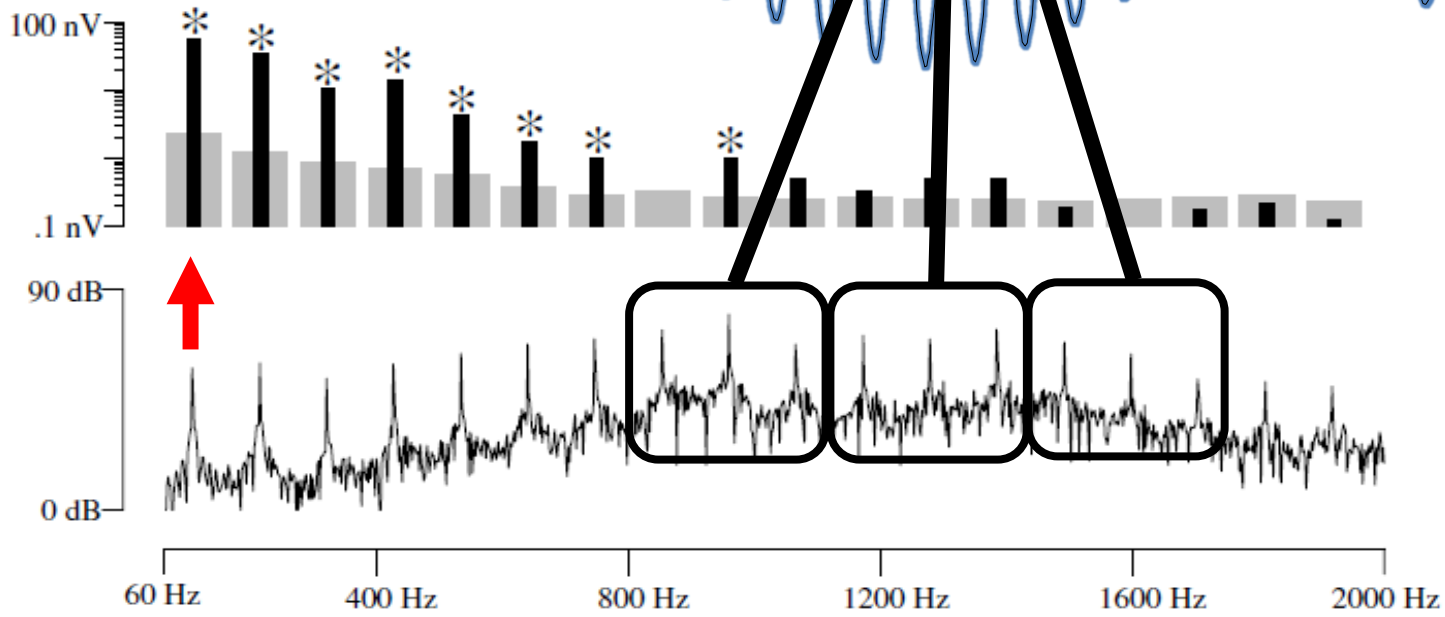
- harmonics are all multiples of this frequency (e.g., 100, 200, 300 etc.)
  - 900 + 1000 + 1100 = fluctuation at 100 Hz
  - 1200 + 1300 + 1400 = fluctuation at 100 Hz
  - 1500 + 1600 + 1700 = fluctuation at 100 Hz





EFR/ASSR

an /a/ vowel



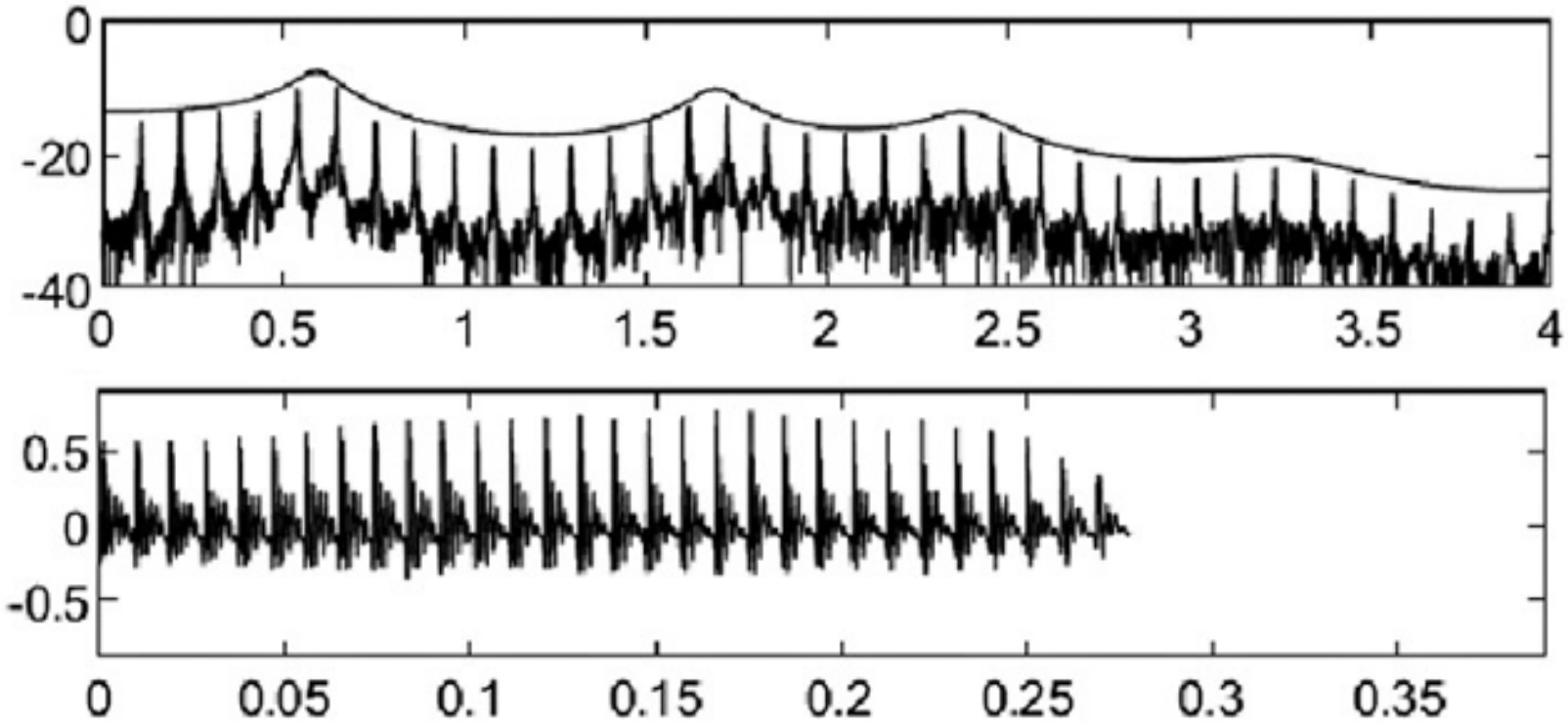
Aiken & Picton, Hear Res, 2008

What does this look like with real speech?

- below is a an /ε/ vowel (as in head)
  - evenly spaced harmonics in top panel (spectrum)
  - these combine to give you a huge fluctuation at fundamental frequency

spectrum

waveform



Easwar et al. Sem Hearing, In Press

Tying it Up

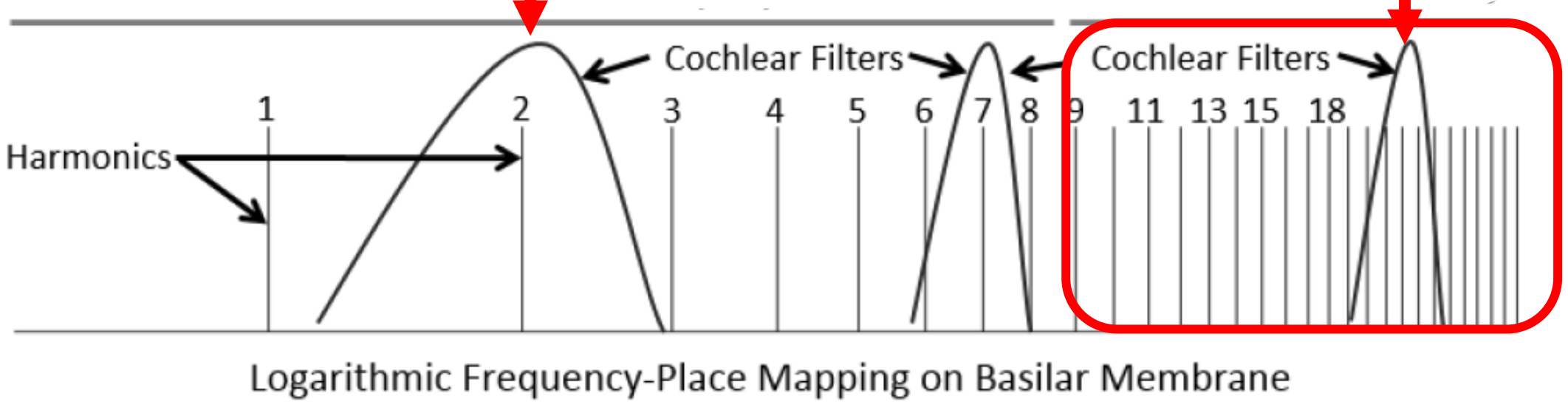
- clinical speech responses are **always** envelope following responses (ASSR)
  - they use alternating polarity
  - they are largest at the speech fundamental frequency
  - they follow a large fluctuation at the speech fundamental frequency
  - this fluctuation *comes from speech harmonics adding together in the cochlea*

Getting back to our question!

Why do we always use male speakers to evoke speech responses?

What Happens to Speech in the Auditory System?

- speech harmonics are *evenly* spaced in frequency, but *logarithmically* spaced in the cochlea
- low harmonics are far from each other;
- only the close harmonics can add together in the cochlea to give us a large fluctuation
  - the big response at the fundamental comes from the high frequency part of speech!



Logarithmic Frequency-Place Mapping on Basilar Membrane

What Happens to Speech in the Auditory System?

- male voices tend to have lower fundamental frequencies, which means the harmonics are closer together (and can add more easily)
  - a fundamental frequency of 100 Hz means all harmonics are spaced by 100 Hz
  - a fundamental frequency of 200 Hz means all harmonics are spaced by 200 Hz

- **Is this the only reason?**
- **Has anyone checked?**

We Investigated This!

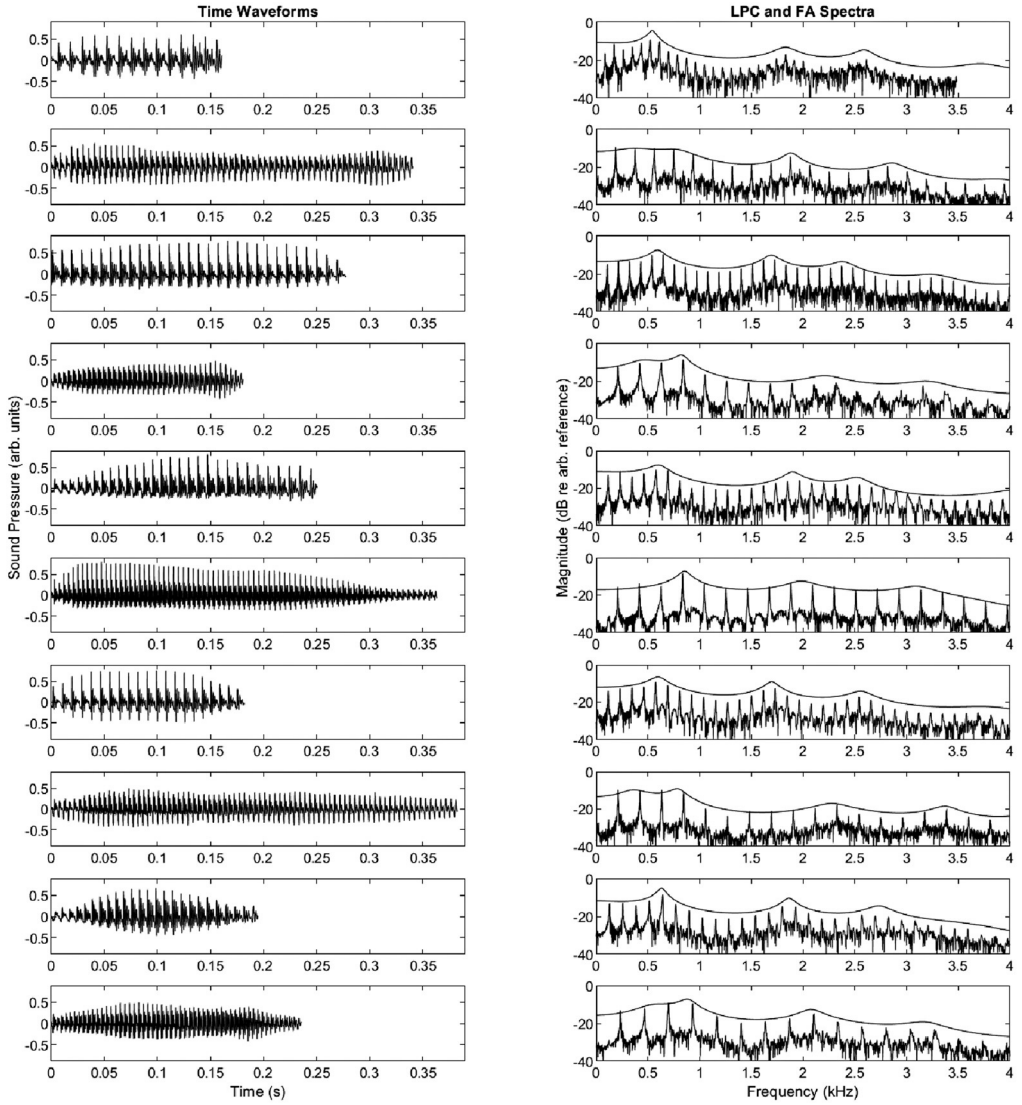
The Influence of Male- and Female-Spoken Vowel Acoustics on Envelope-Following Responses

Vijayalakshmi Easwar[Q1], Ph.D.,[1] David Purcell, Ph.D.,[2]
Maaike Van Eeckhoutte, Ph.D.,[3,4,5] and Steven J. Aiken, Ph.D.[6]

- Questions
  - How can we account the differences between male and female voices?
  - What are the acoustic factors that make responses smaller to female voices?

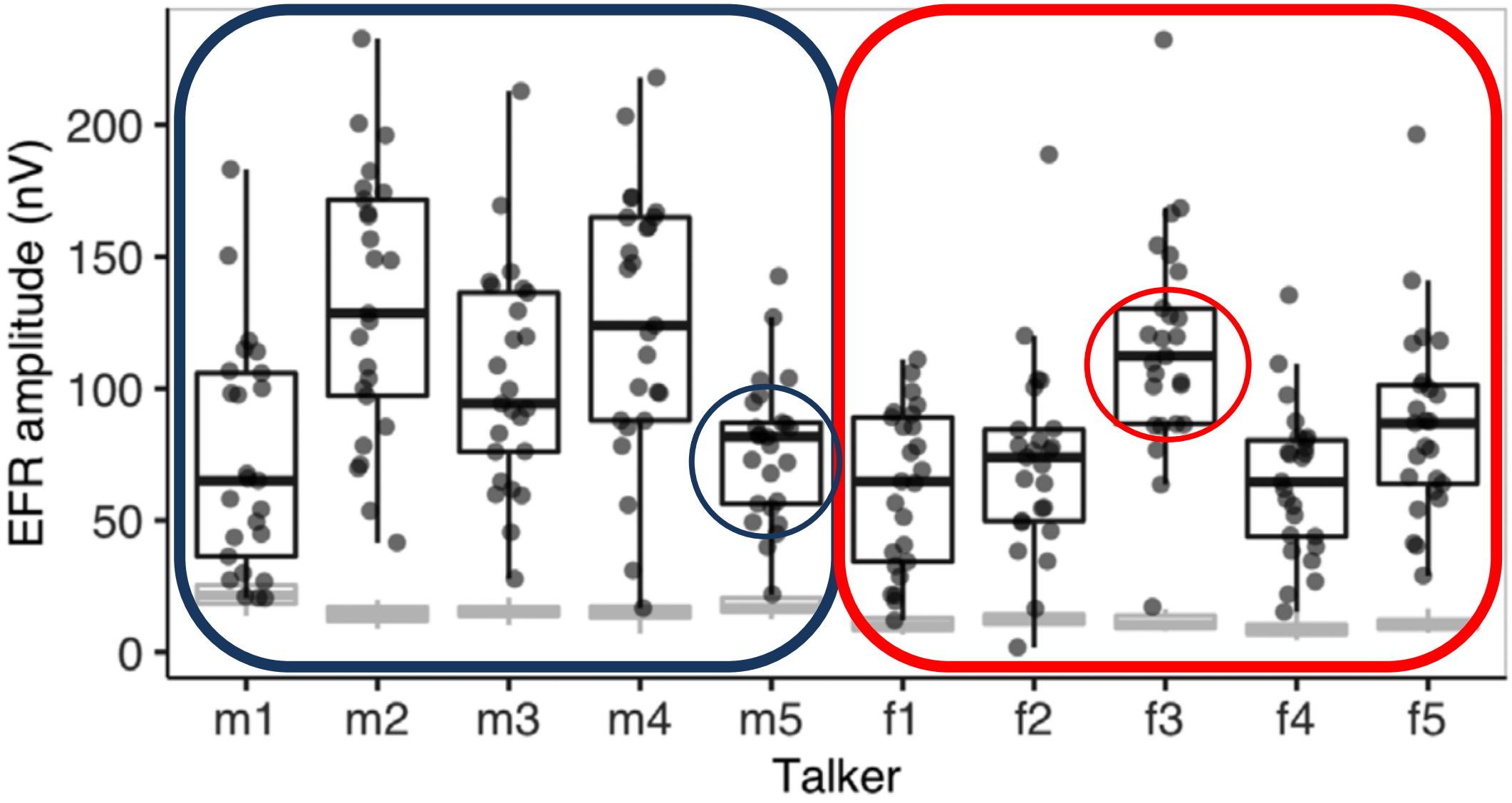
- Knowing this is important, if we want to use these responses to study hearing (and not just for male voices)

Study Design

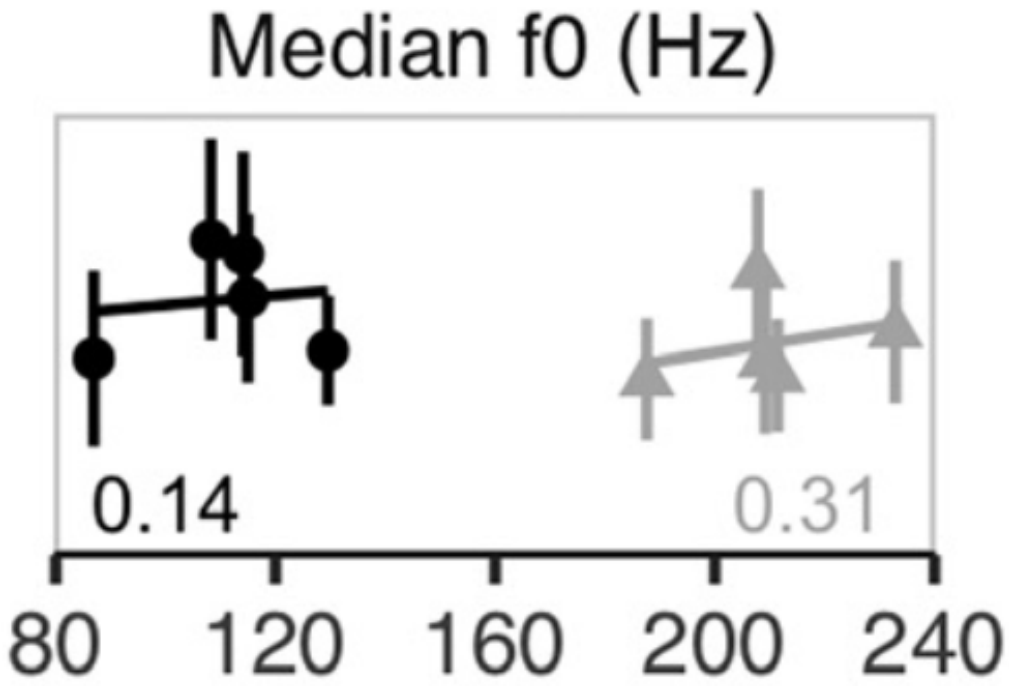


- 10 speakers saying /ɛ/ as in head
  - SPEAKERS: **5 male** and **5 female**
  - LISTENERS: 25 adults (21 female) with normal hearing
- participants listened to all 10 versions in each polarity, interleaved

Results: EFR Amplitude for All Talkers



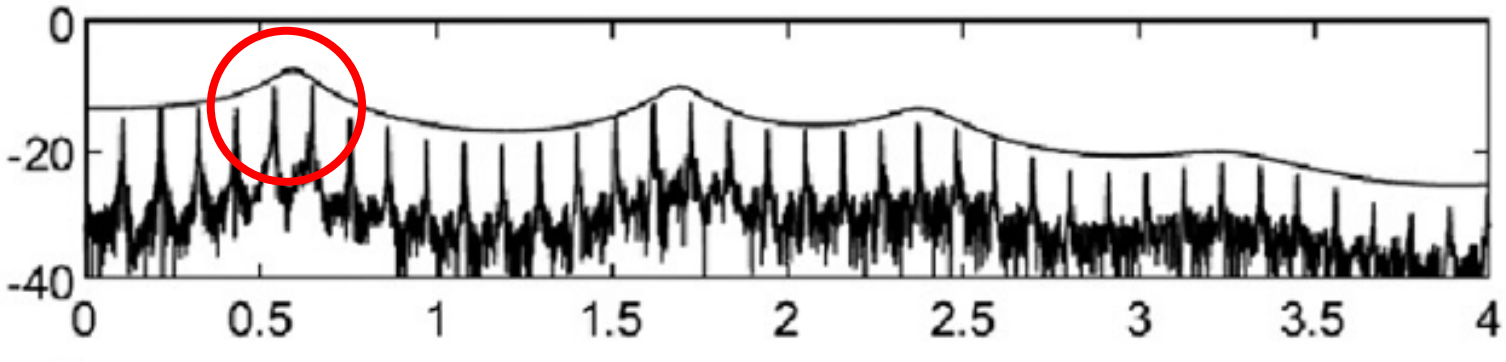
Was it mostly because of fundamental frequency?

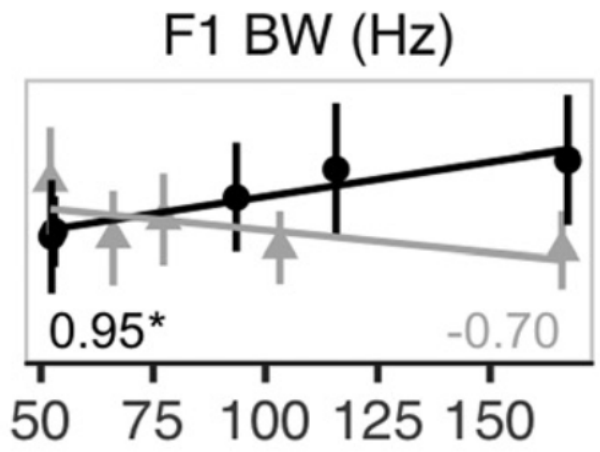


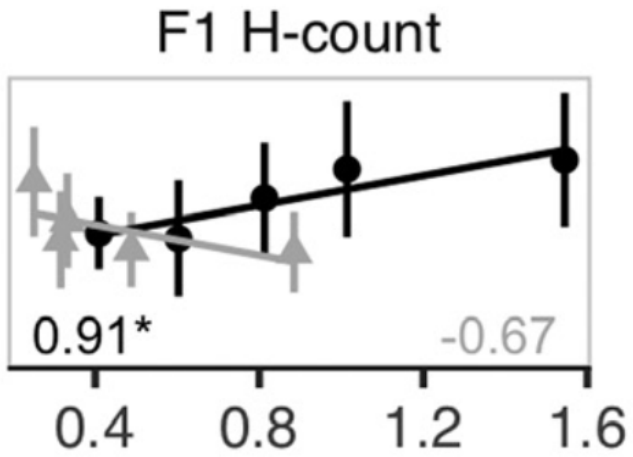
- fundamental frequency did **not** correlate strongly with response amplitude
- the correlation within each group was positive, not negative!

What acoustic features **actually** mattered?

- for male speakers, **two things** were correlated with response amplitude
  1. bandwidth of the first formant
  2. number of harmonics in the first formant

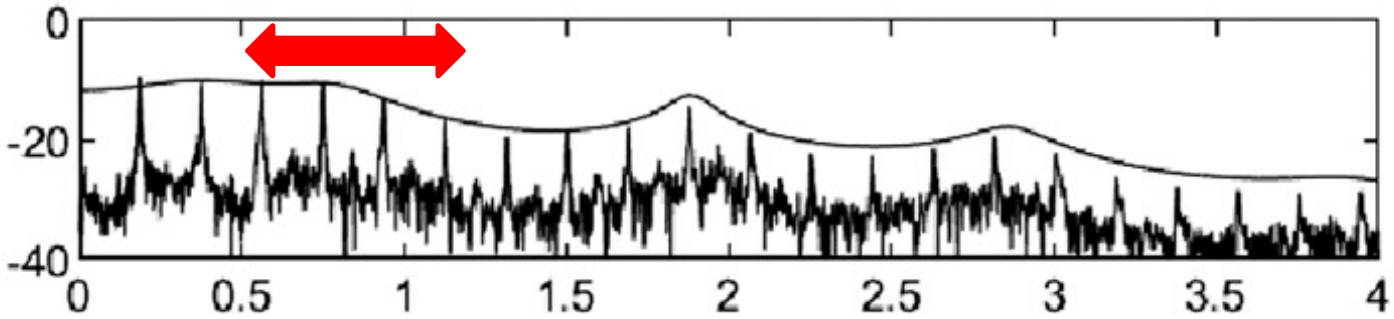


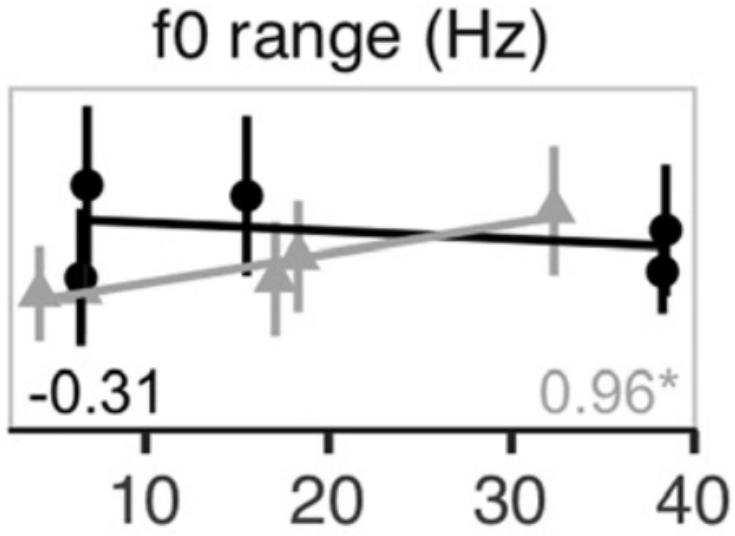




What acoustic features **actually** mattered?

- for female speakers, **one thing** was correlated with response amplitude
  - the fundamental frequency range (more variable pitches were better)





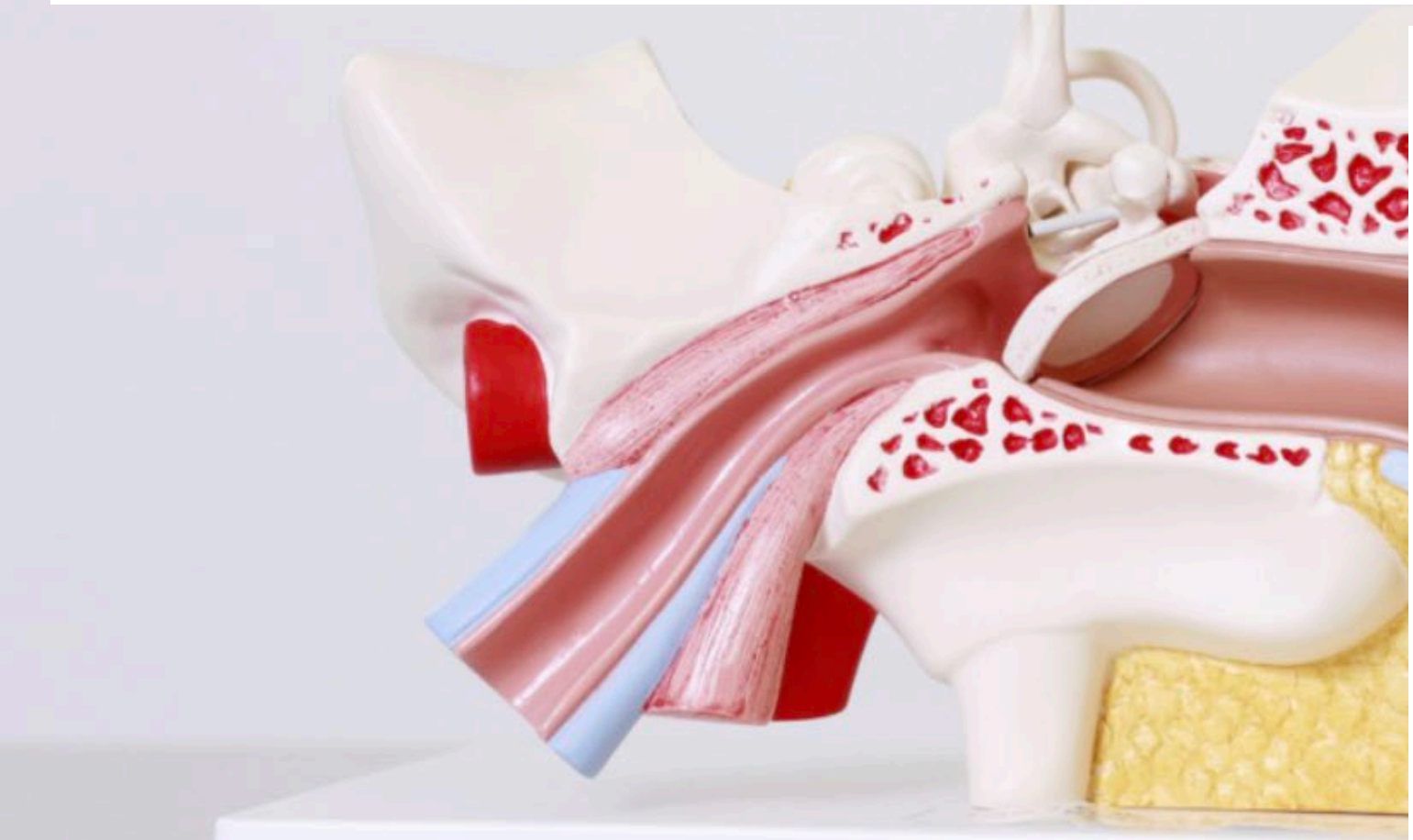
Why are male voices better?

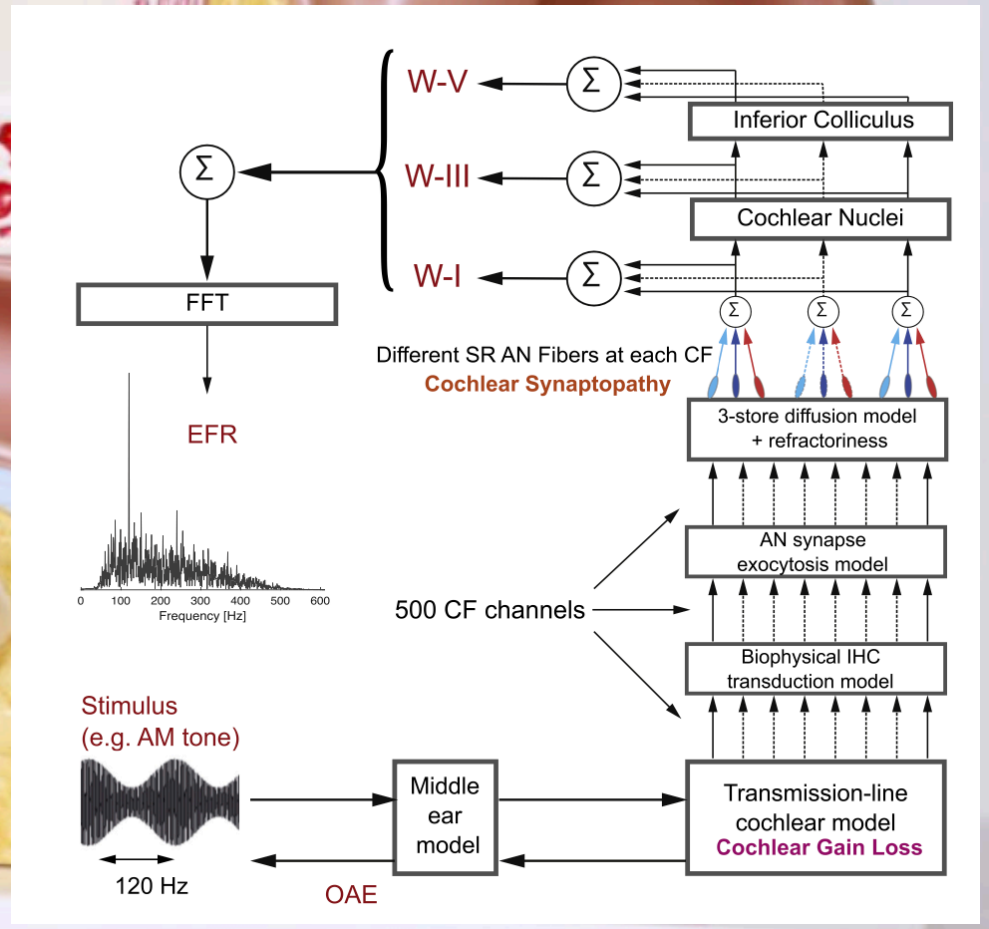
- Well actually… they're **not** always better!
  - our third female speaker gave larger responses than two of our male speakers
- However, the reasons for the differences were still hard to understand
  - the first formant appears to be important for male speakers
  - only pitch range had any relationship for female speakers
- Can we get a clearer picture of what is going on?

Ear Modelling

Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss

Sarah Verhulst [a,*], Alessandro Altoè [b], Viacheslav Vasilkov [c]



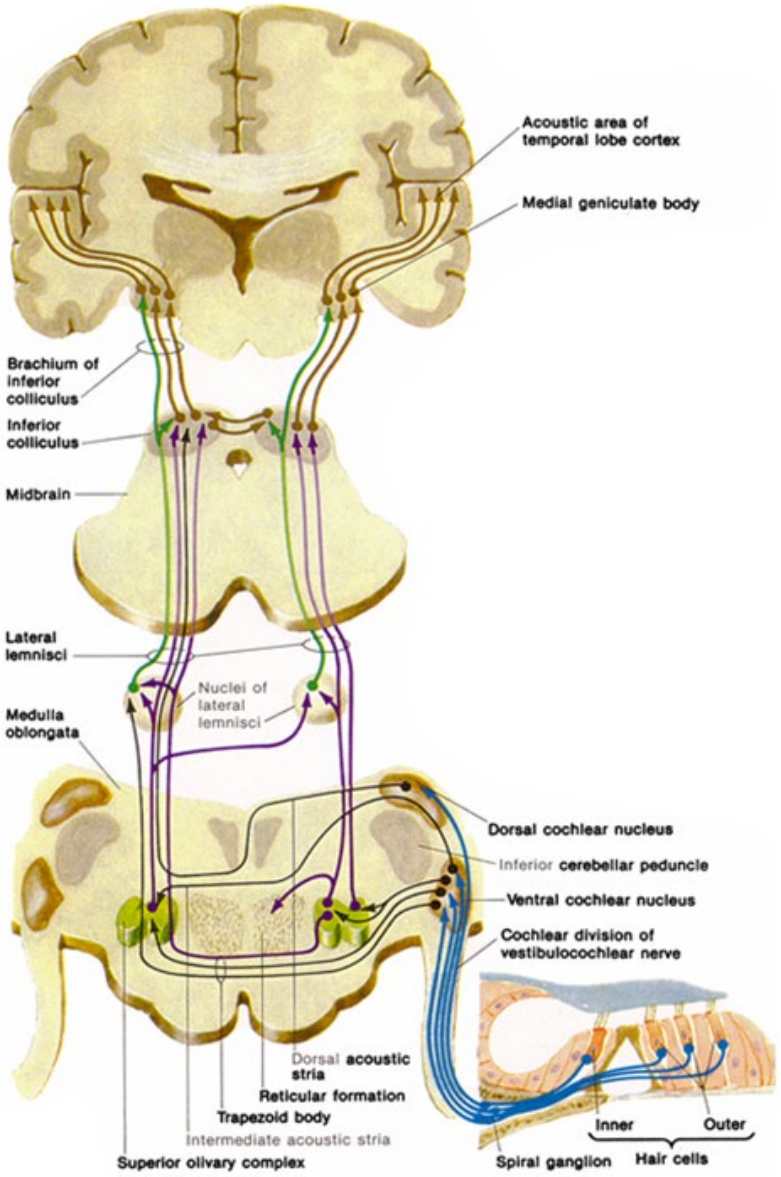


Ear Modelling


Kate McDowell, Porcelain

- a mathematical model that estimates how the auditory system processes sounds
  - middle ear
  - cochlea
  - auditory nerve
  - cochlear nucleus
  - inferior colliculus
  - waves I, III and V of the ABR
  - the envelope-following response

Ear Modelling

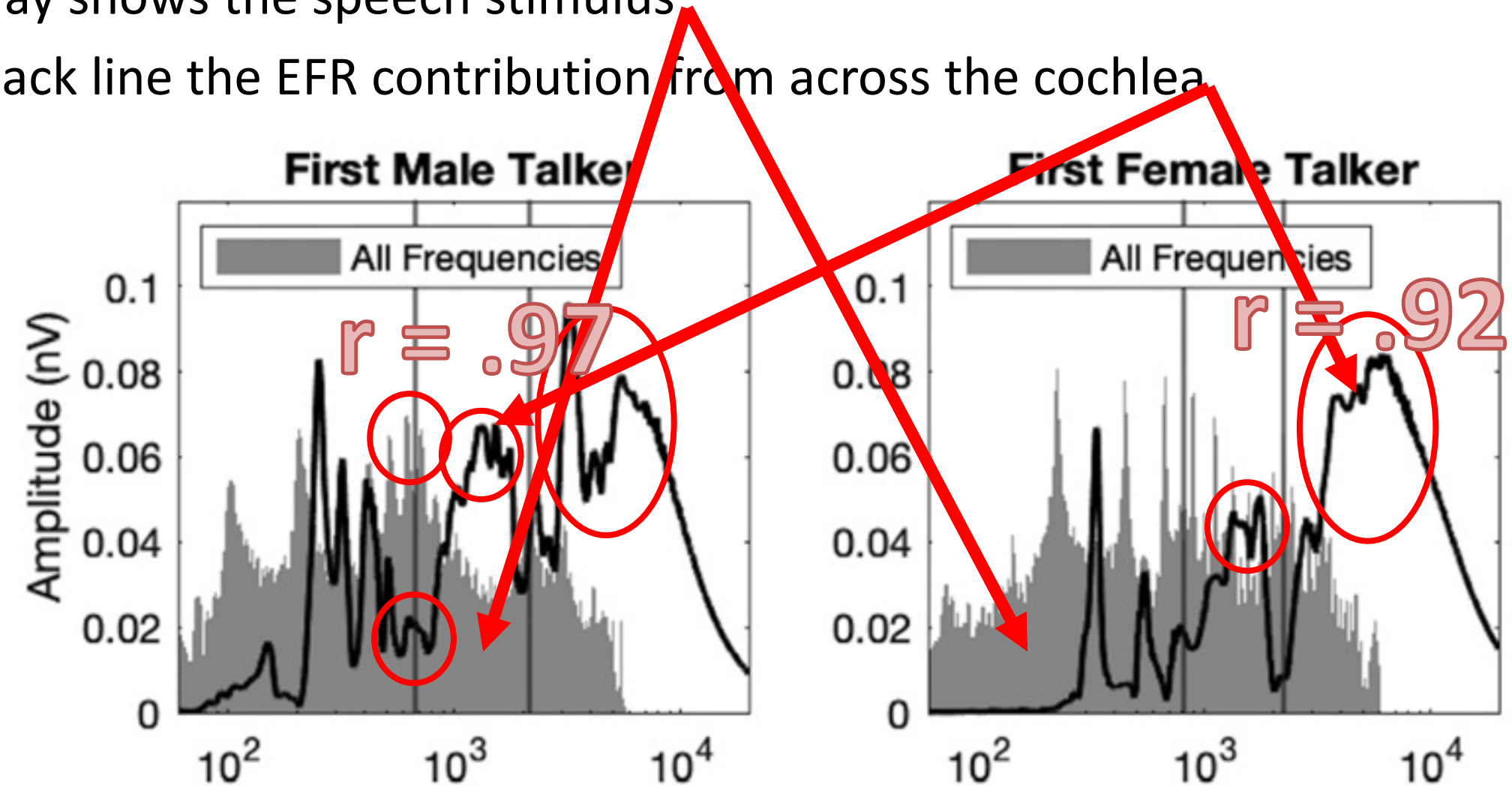


- the envelope-following response comes from the auditory nerve and brainstem—coming from the **whole cochlea**

- the model made it possible to see how much of the response depended on **each part of the cochlea**

Stimulus and Envelope Response Across the Cochlea

- gray shows the speech stimulus
- black line the EFR contribution from across the cochlea



Ear Modelling Observations


Kate McDowell, Porcelain

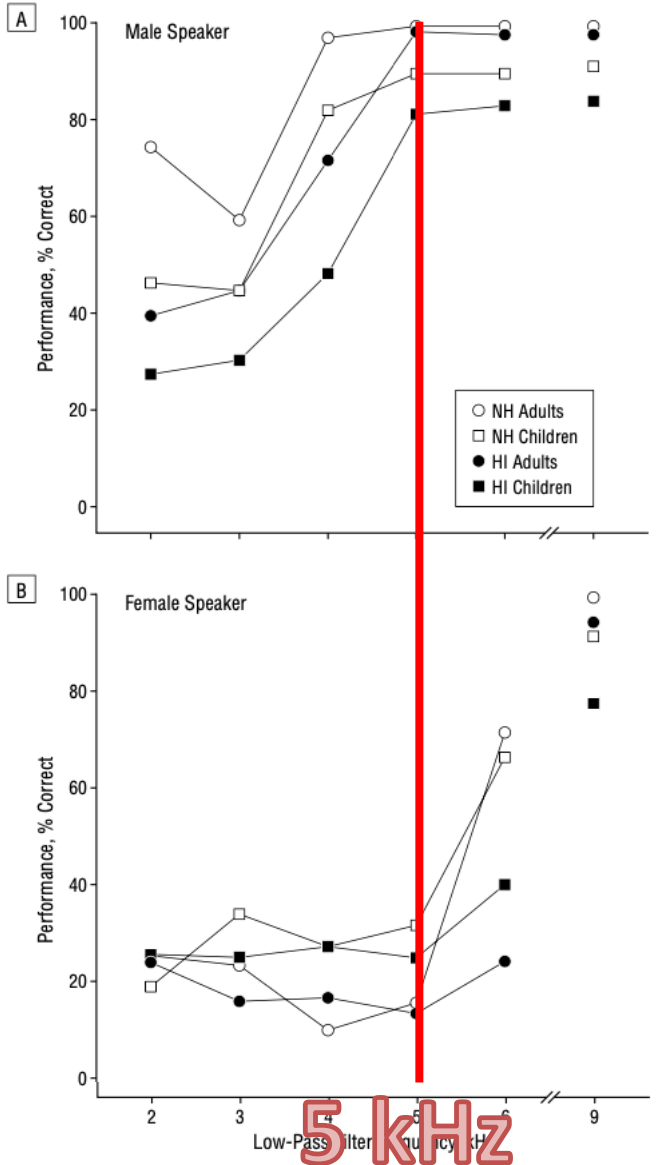
- **male** voices → response predicted by activity above the **first** formant
- female voices → **response** predicted activity above the **second** formant
- so why do responses tend to be smaller for female speakers?
  - speech levels are much softer above the second formant (6 dB/octave)

Can we do better?

- Almost certainly!
  - boost high-frequencies for female voices
- Should we care?
  - female voices give us clearer information about the brainstem (low fundamental frequencies may have contributions from the cortex, see Coffey et al., 2016)
  - female voices are important
- Does it matter for hearing?

Does this Matter for Hearing?

- the study showed that male and female voices can be quite different
  - female voices *may* need slightly more gain in the high frequencies (near F2)
  - female voices *do* benefit from wider high-frequency bandwidth for fricative perception (Stelmachowicz et al., 2004)



Wrapping things up…

- the speech envelope-following response is a large response at the fundamental frequency (voice pitch)

- it primarily reflects encoding of **high** speech harmonics
  - first formant and above for male voices
  - second formant and above for female voices

- male and female voices are different, and this is probably important for hearing

- **research in hearing speech should always include male and female voices**

Including female voices… it's