Converging Hearing and Speech Enhancement Technologies

Fan-Gang Zeng

University of California, Irvine

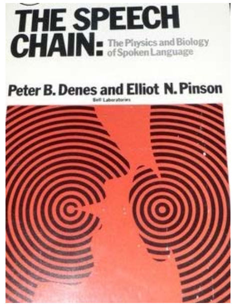
fzeng@uci.edu;

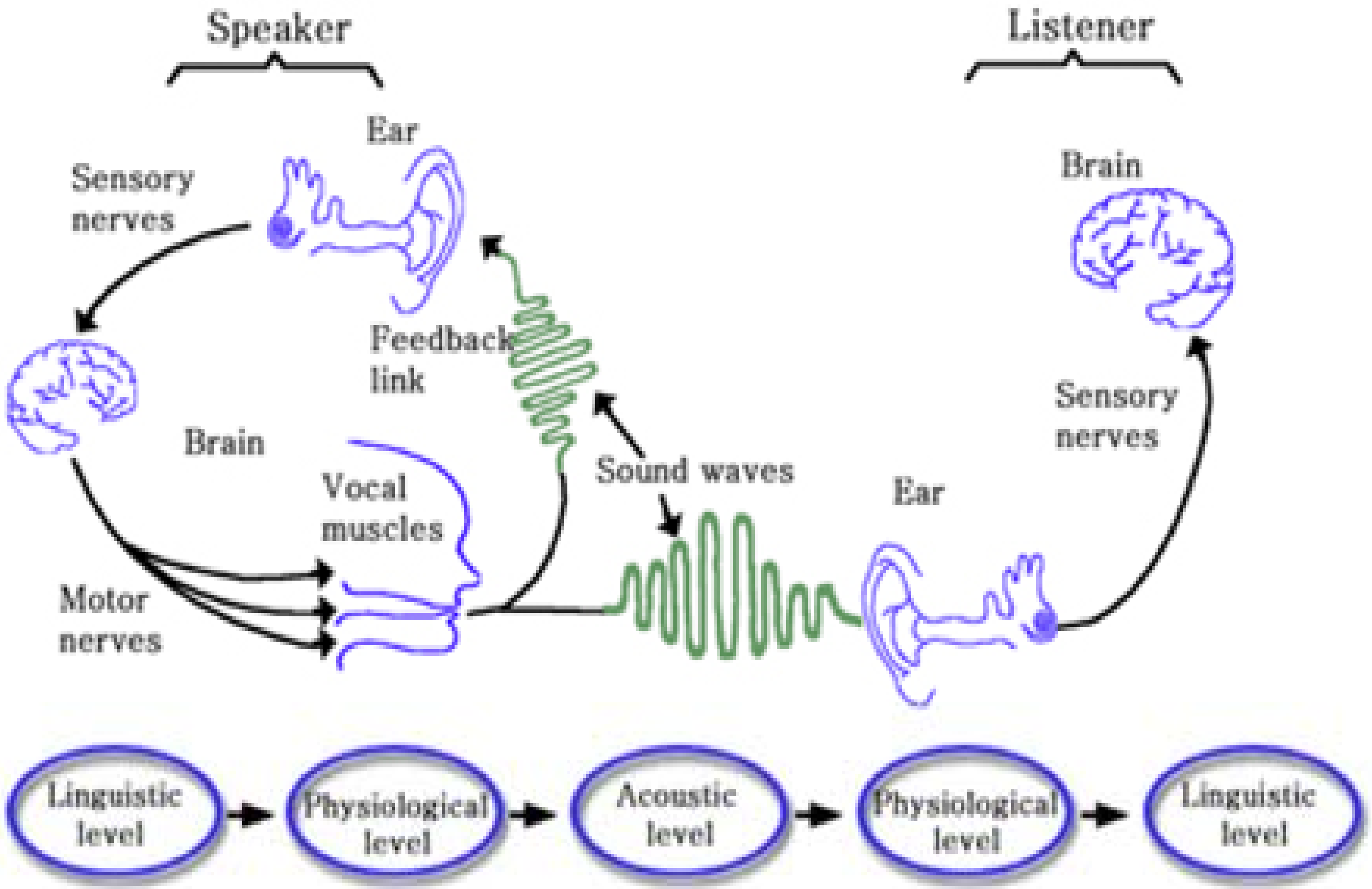
http://www.ucihs.uci.edu/hesp

Contents

The Speech Chain:

From Production to Perception



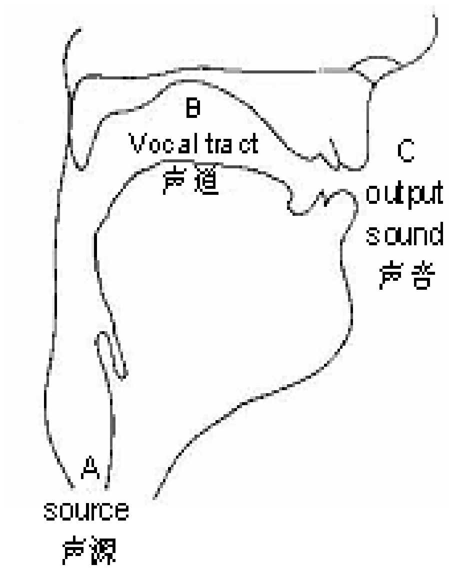


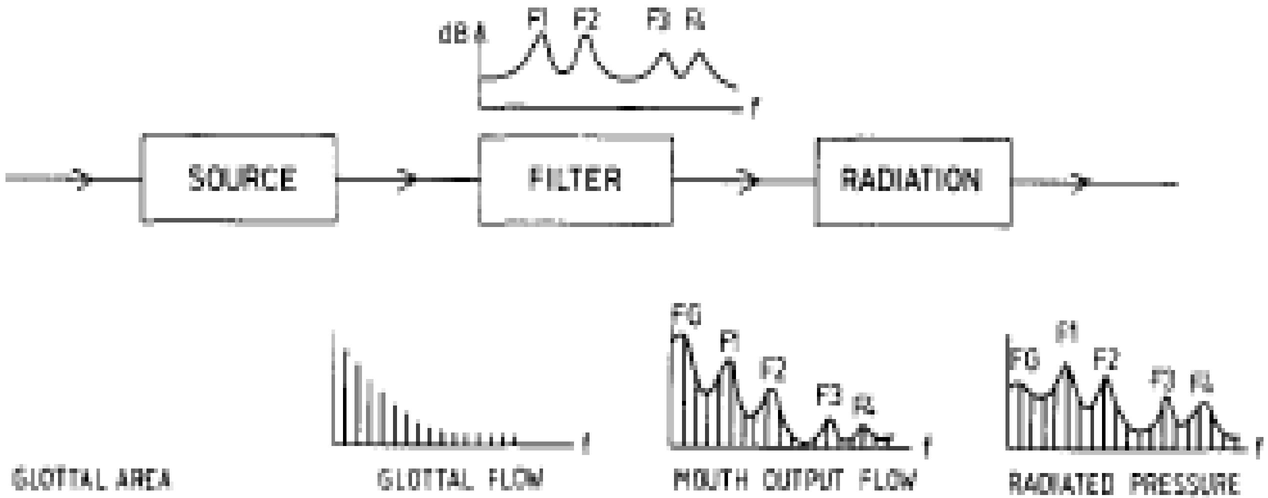
The *broken* speech chain:

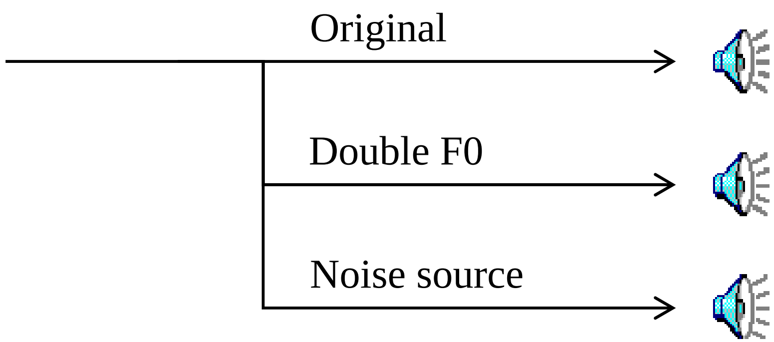
Opportunities for enhancement

Production:

Source : Filter = Fine structure : Envelope





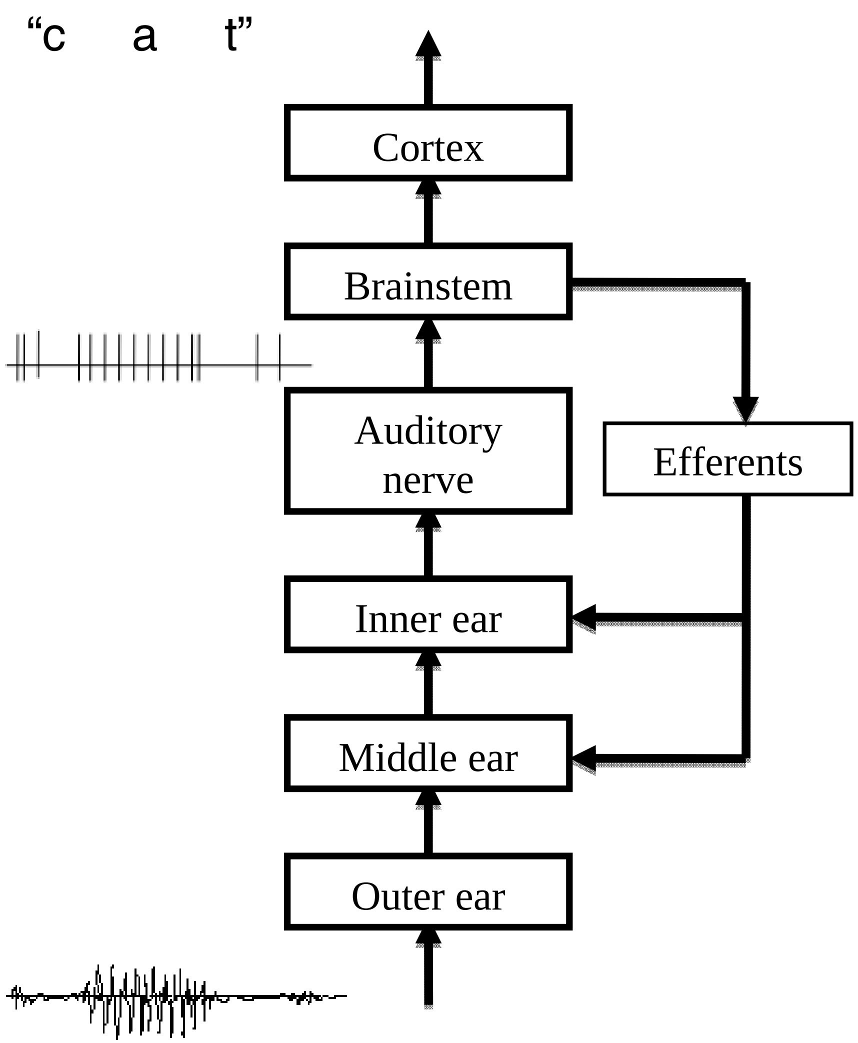


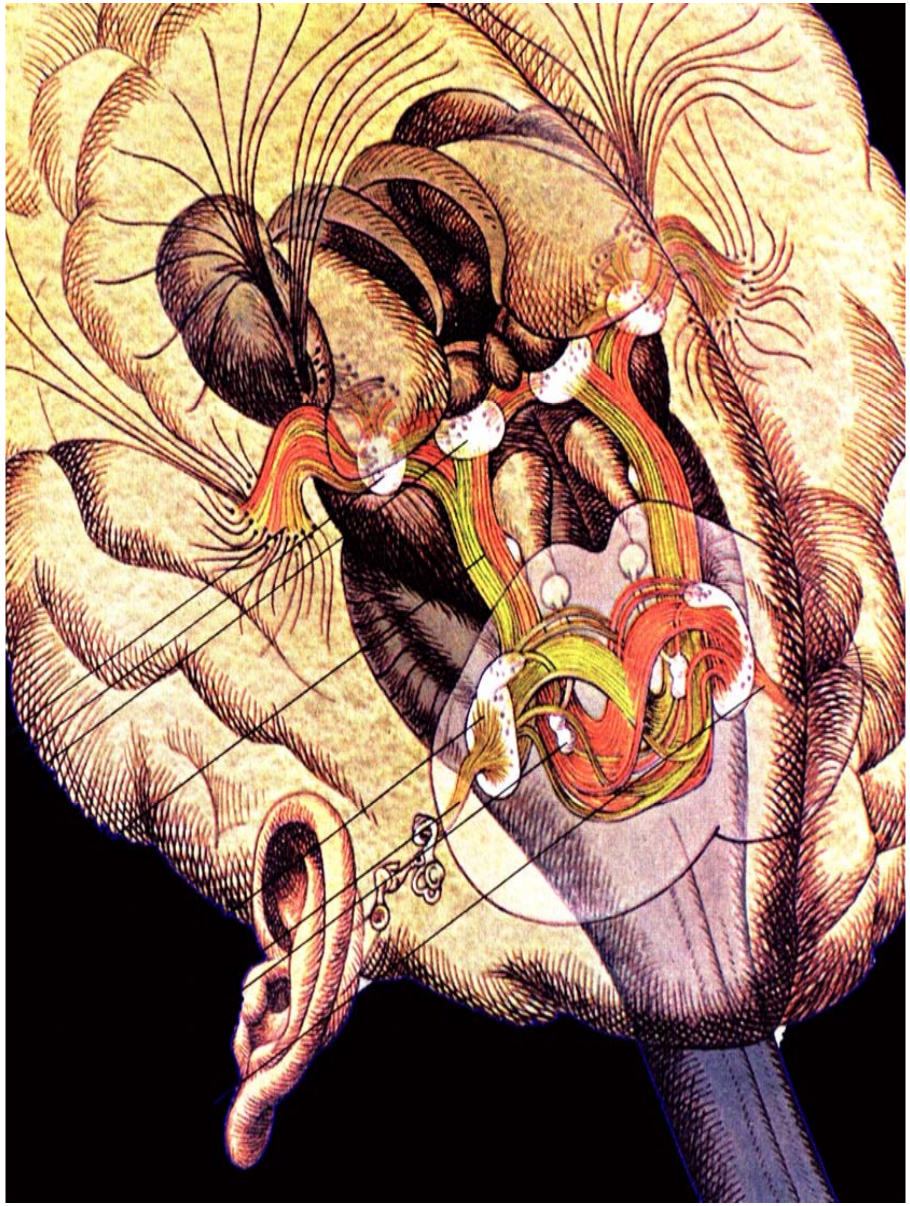
Gunnar Fant (1919-2009)

Kawahara et al. (1999) Speech Communication

The amazing speech

Perception:
Peripheral and central processing



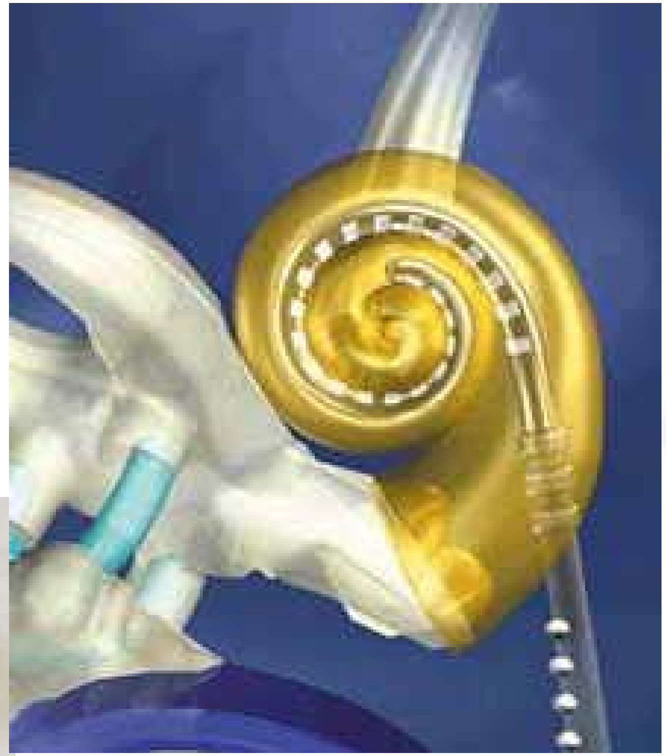


The amazing hearing

You can hear from whisper to 110-dB PA in a moment like this…

Cochlear implant status

- 200,000 users worldwide
- Performance = 80% in quiet
- Talk on the phone
- $1B revenue, $25B cap
- 3 FDA approved devices
- 10 start ups

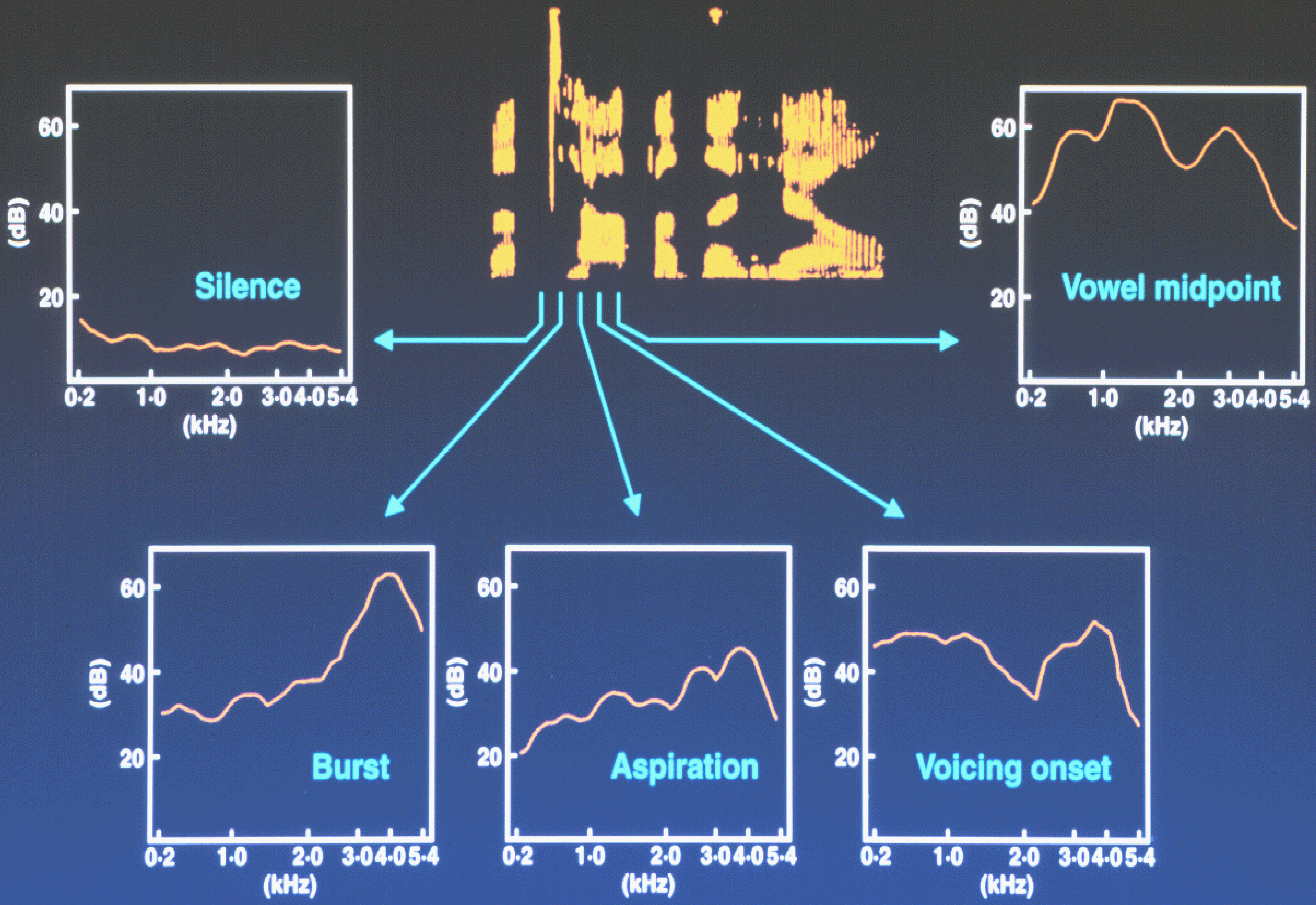


- Power hungry
- Expensive
- Imperfect hearing: CI music vs. Original

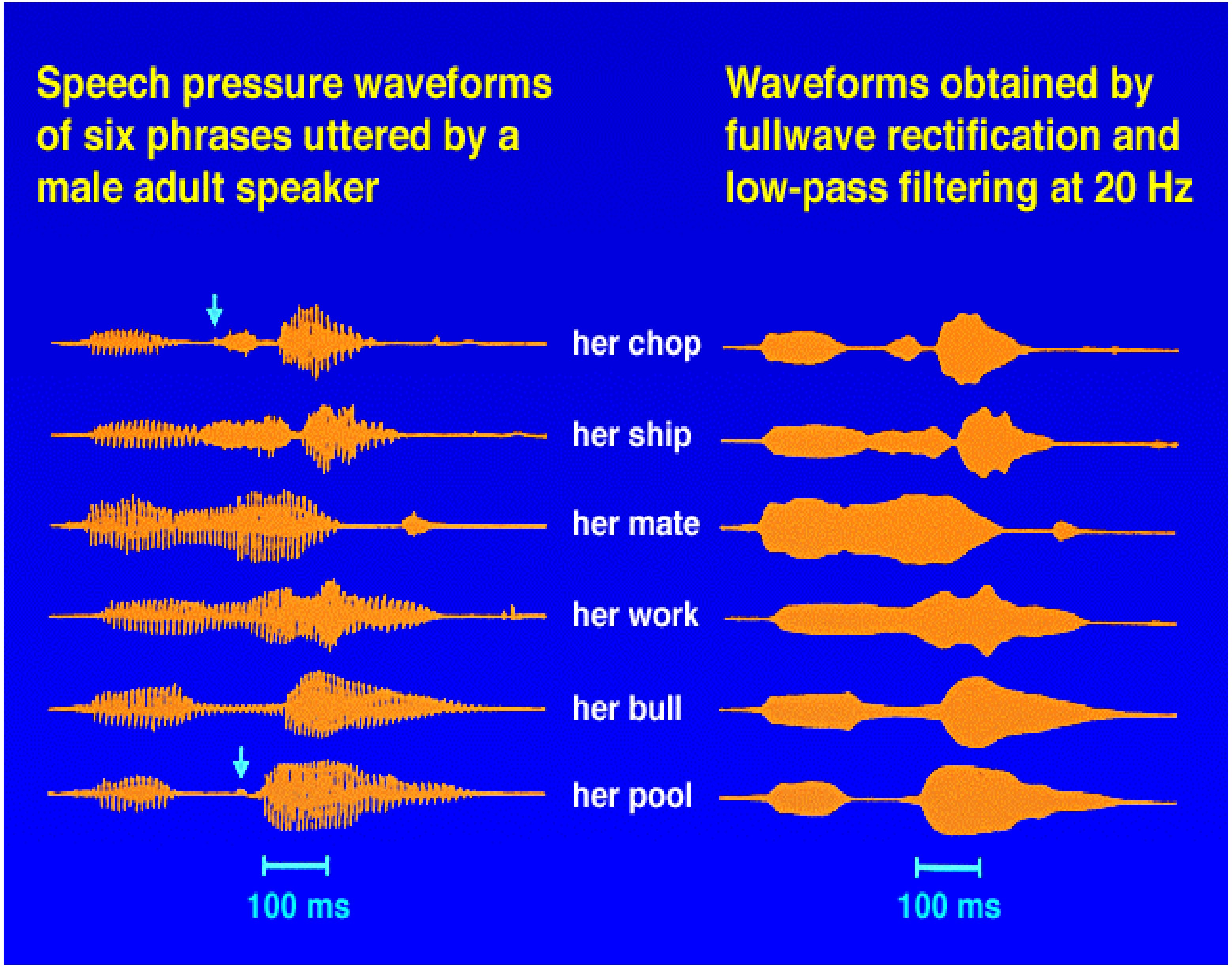
Summary

- Chapter 1: Our work is interesting

Broadband spectrogram of the phrase "the top of the hill"

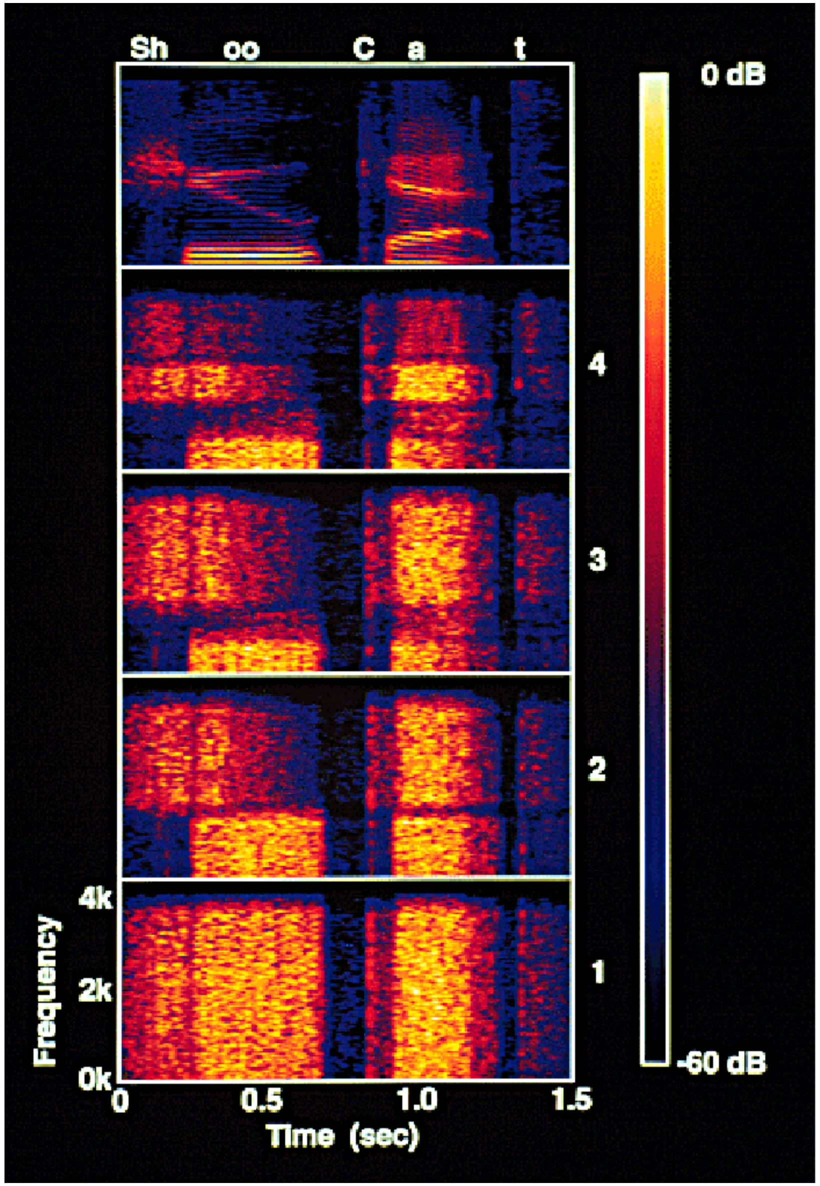


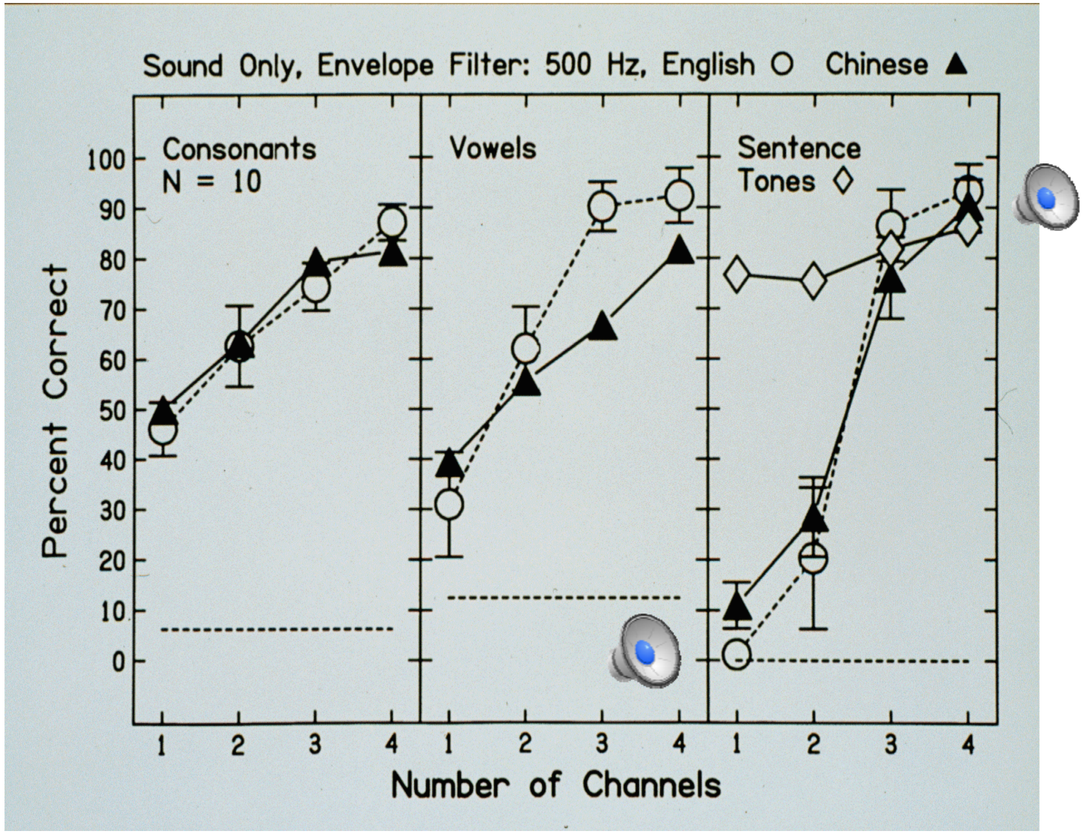
Temporal envelope cues



- Rosen (1991) Royal Phil Soc Trans

Speech recognition with primarily temporal cues





Shannon, Zeng, Kamath et al. (1995) Science

Little math

- Flanagan (1980) “Parametric coding of speech spectra”

$$s(t) \approx \sum_{k=1}^{N} A_k(t)\cos\left[2\pi f_{ck} t + 2\pi \int_0^t \dot{\phi}_k(\tau)d\tau + \theta_k\right]$$

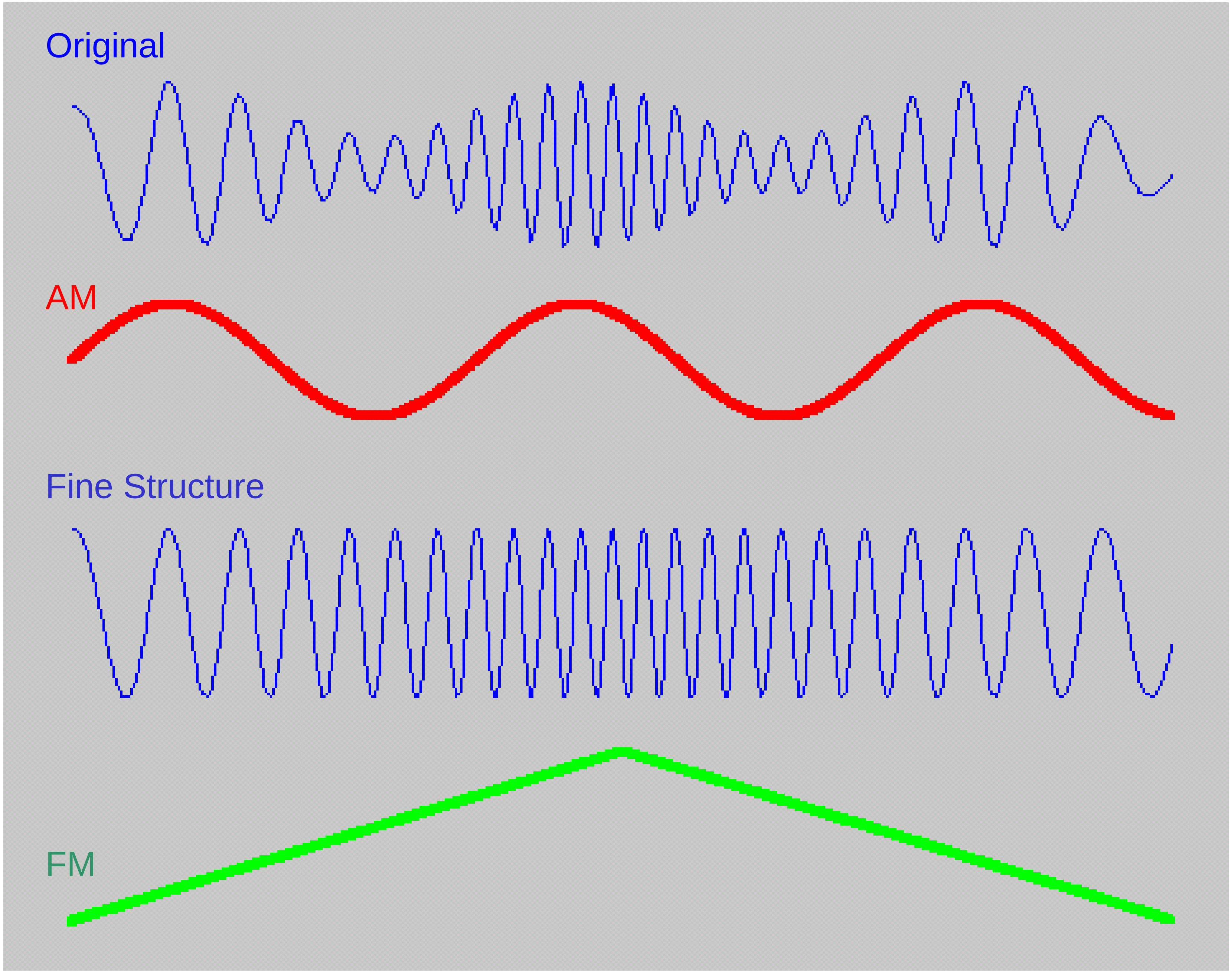
  – Discard absolute phase:

$$s(t) \approx \sum_{k=1}^{N} A_k(t)\cos\left[2\pi f_{ck} t + 2\pi \int_0^t \dot{\phi}_k(\tau)d\tau\right]$$

  – Discard relative phase (i.e., frequency modulation):

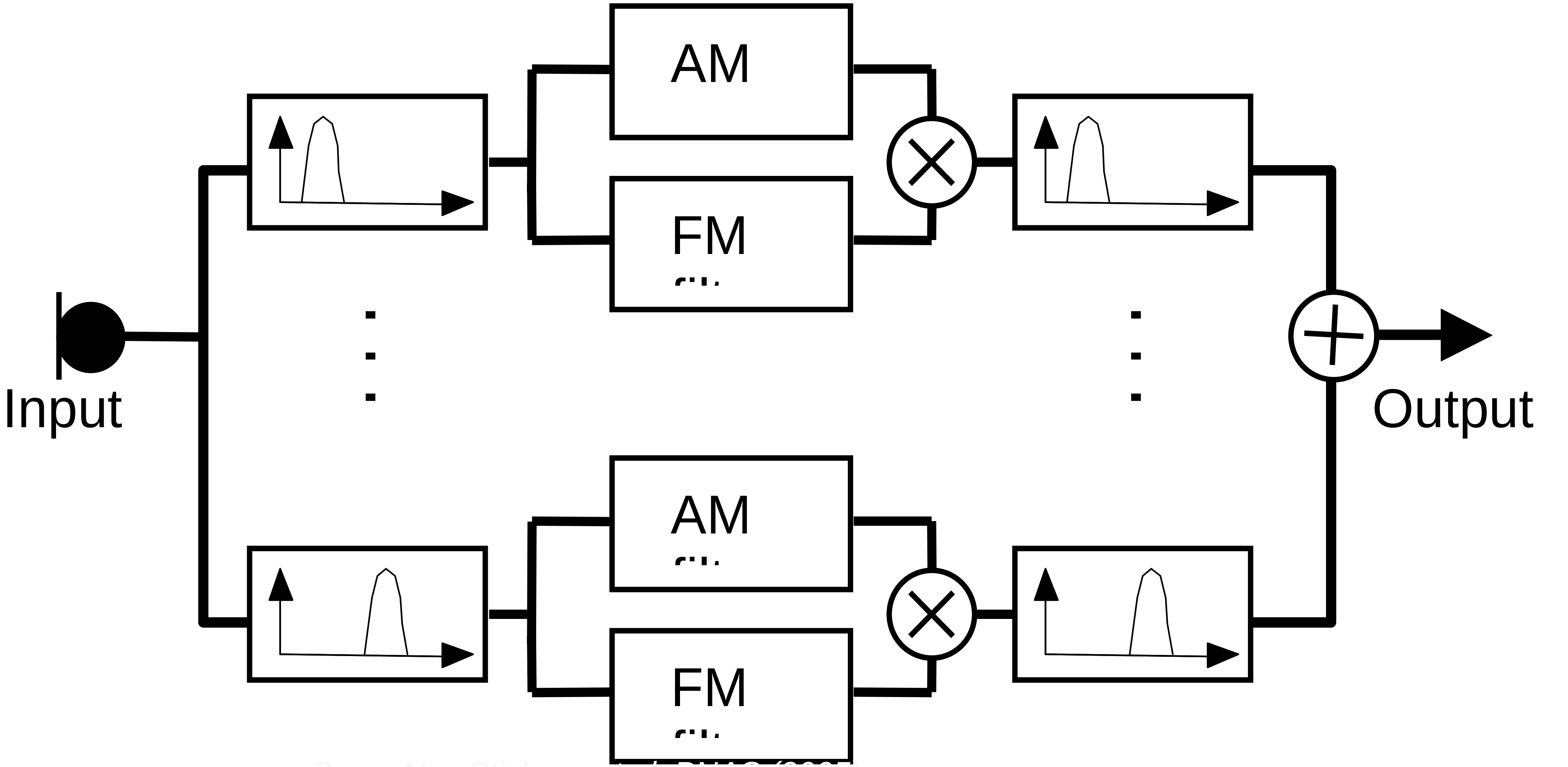
$$s(t) \approx \sum_{k=1}^{N} A_k(t)\cos\left[2\pi f_{ck} t\right]$$

What is fine structure?



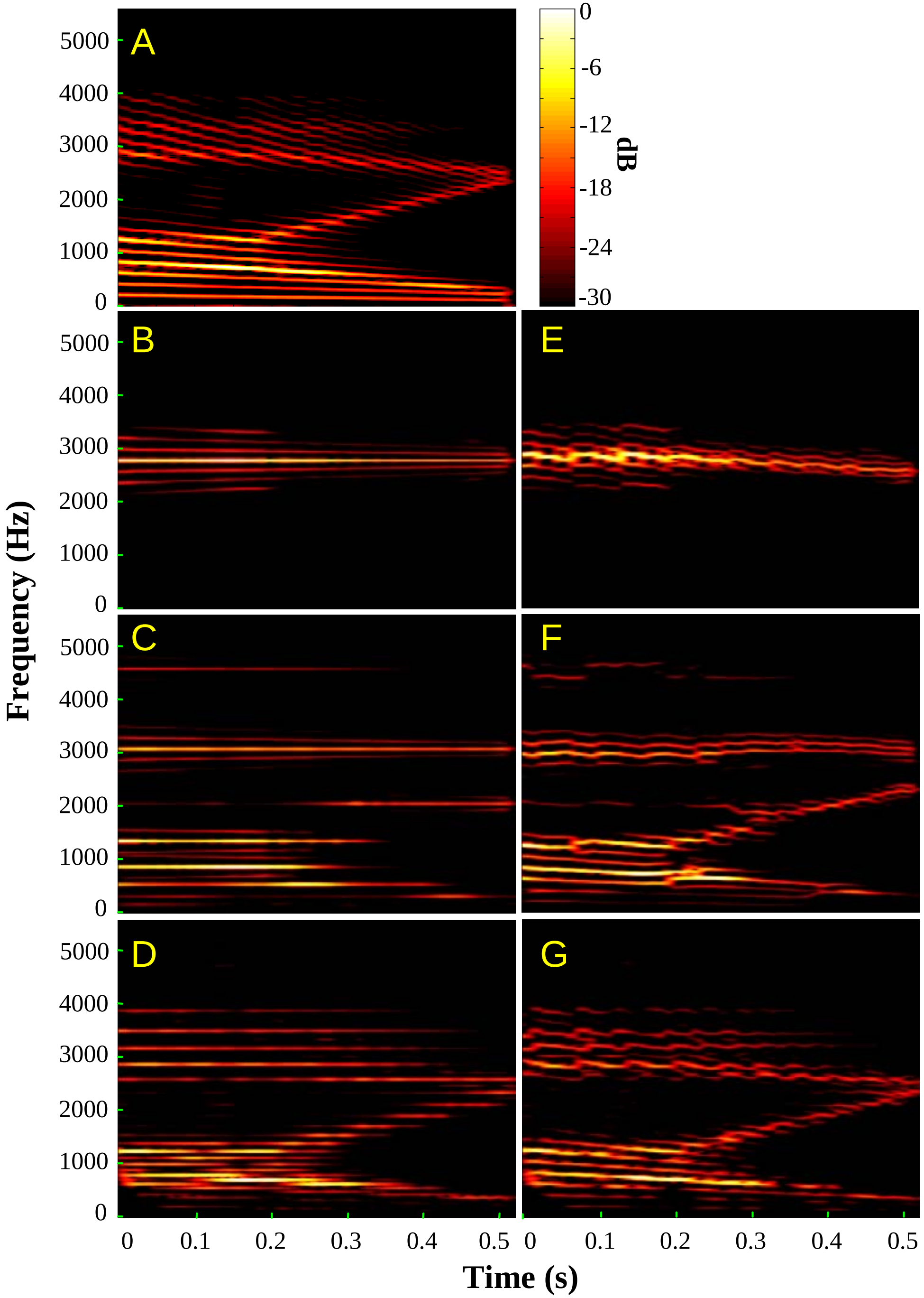
Implementation

- Combo of Dudley’s vocoder and Flanagan’s phase vocoder



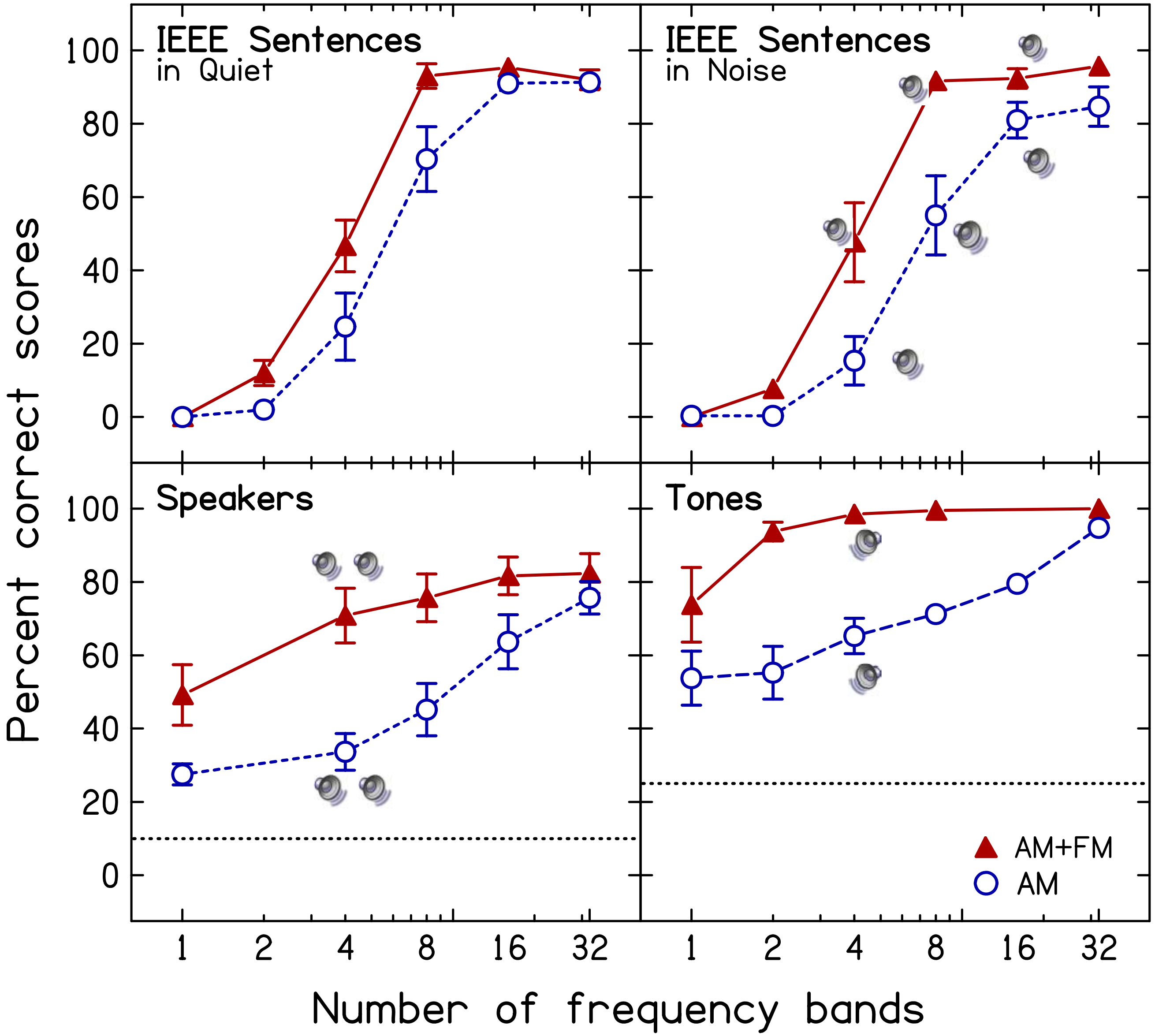
Zeng, Nie, Stickney et al. PNAS (2005)

Spectra: What does FM encode?



Zeng, Nie, Stickney et al.
PNAS (2005)

Sentence, speaker, and tone recognition



Combo:

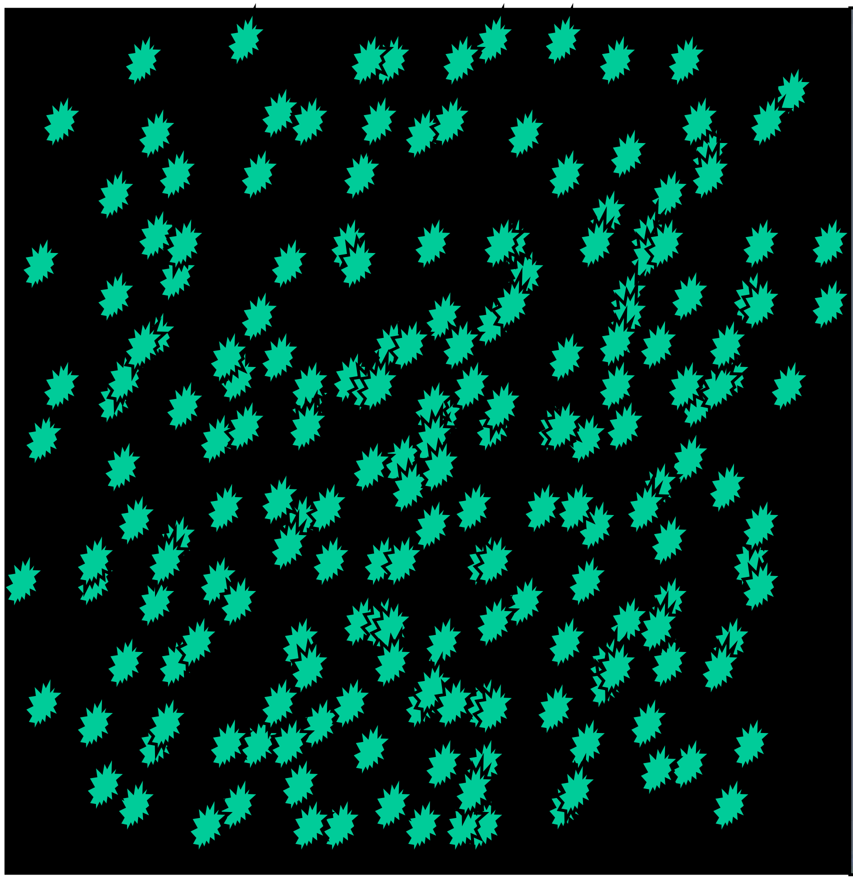
Target:

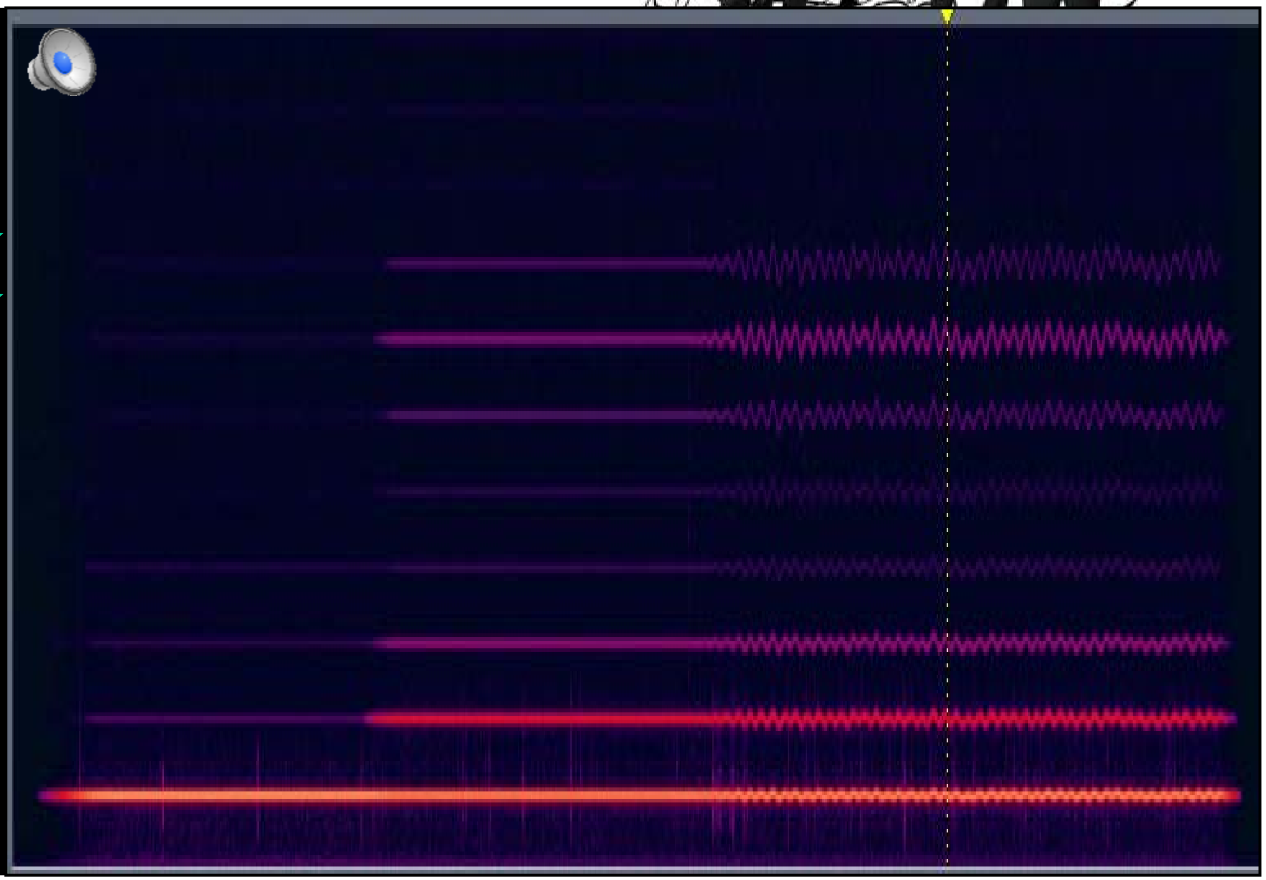
Masker:

Zeng, Nie, Stickney et al. PNAS (2005)

Zeng, Nie, Stickney et al. PNAS (2005)

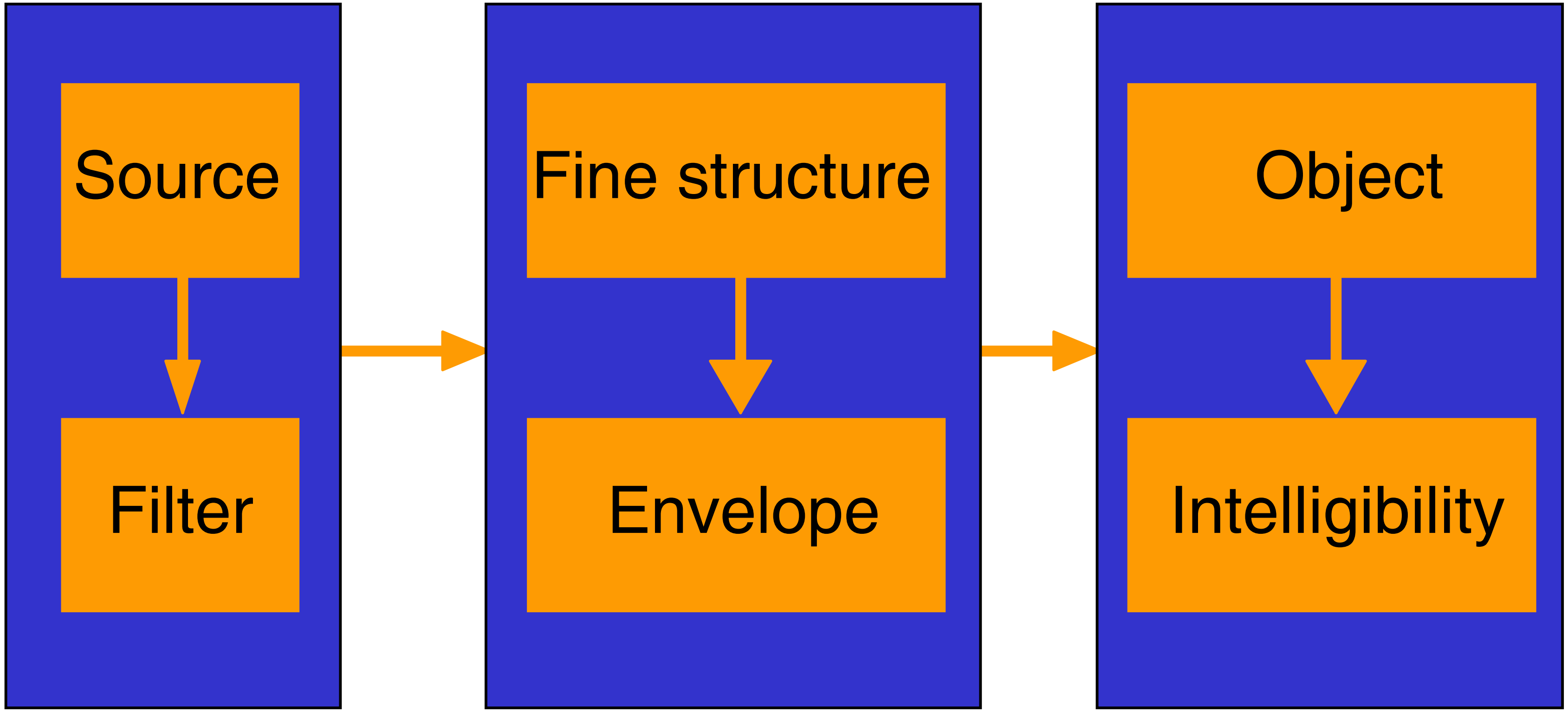
Role of common FM: Binding and segregation





Courtesy of John Chowning at Center for Music Research, Stanford University

A Unified Model

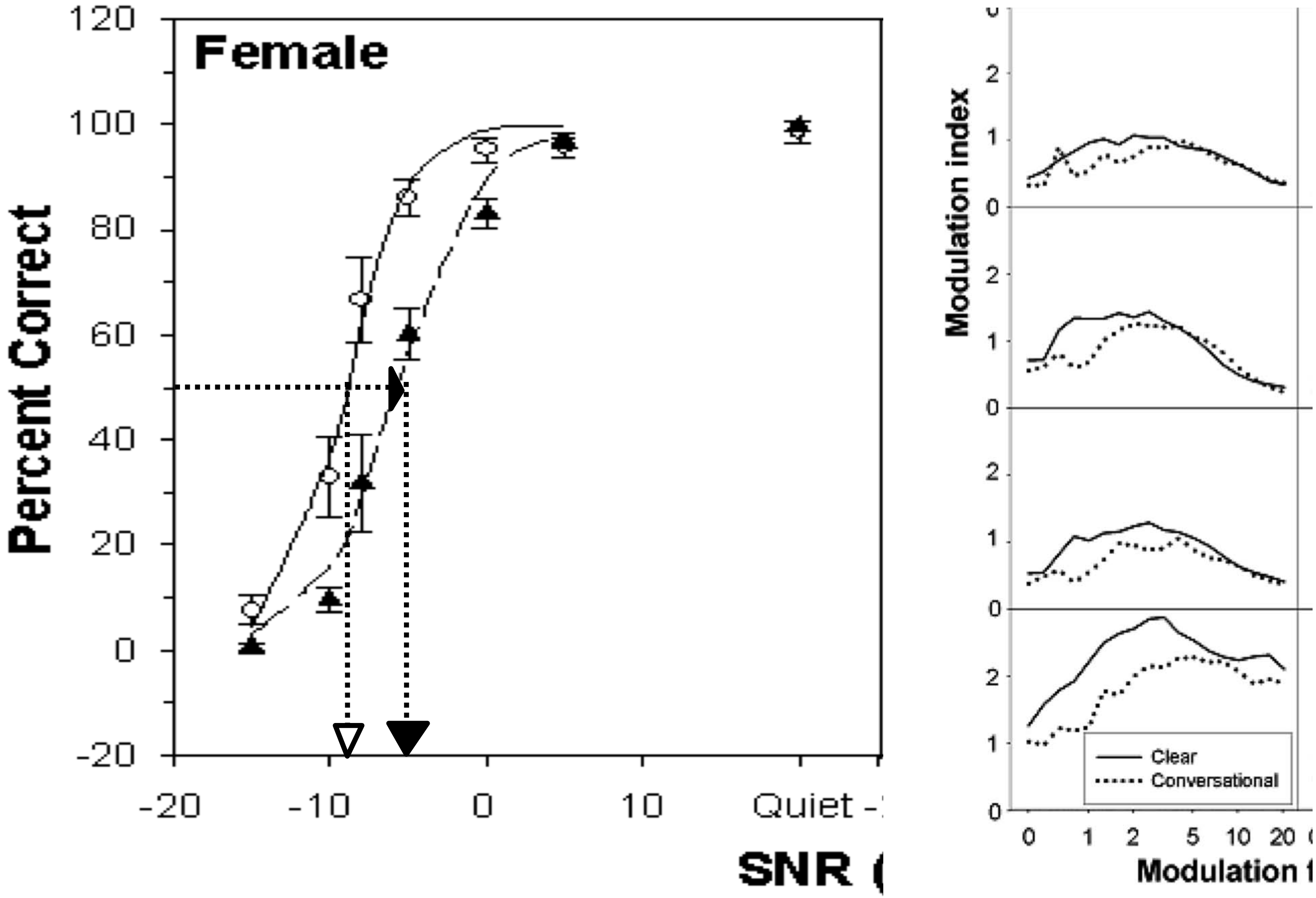


Production Acoustics Perception

Summary

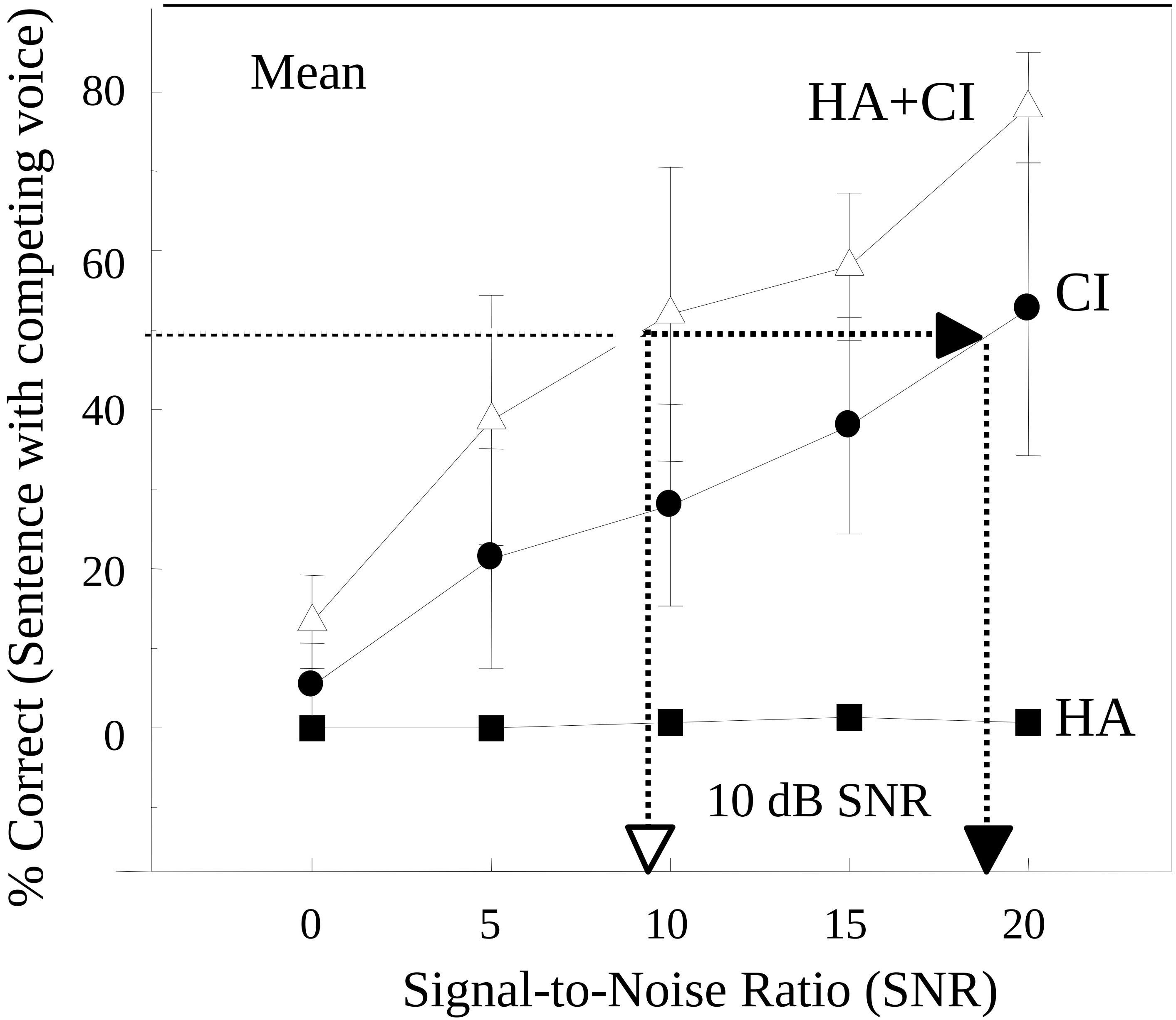
- Chapter 1: Our work is interesting
- Chapter 2: Speech cues are redundant and complementary

Clear speech perception



Liu, Del Rio, Bradlow and Zeng, JASA (2005)

Speech recognition with hearing aid and cochlear implant



Kong, Stickney, and Zeng JASA (2005)

Kong, Stickney and Zeng, JASA (2005)

Summary

- Chapter 1: Our work is interesting
- Chapter 2: Speech cues are redundant and complementary
- Chapter 3: Increasing functional signal-to-noise ratio is the key

Converging technologies:
Hearing aid or Bluetooth headset

www.soundid.com

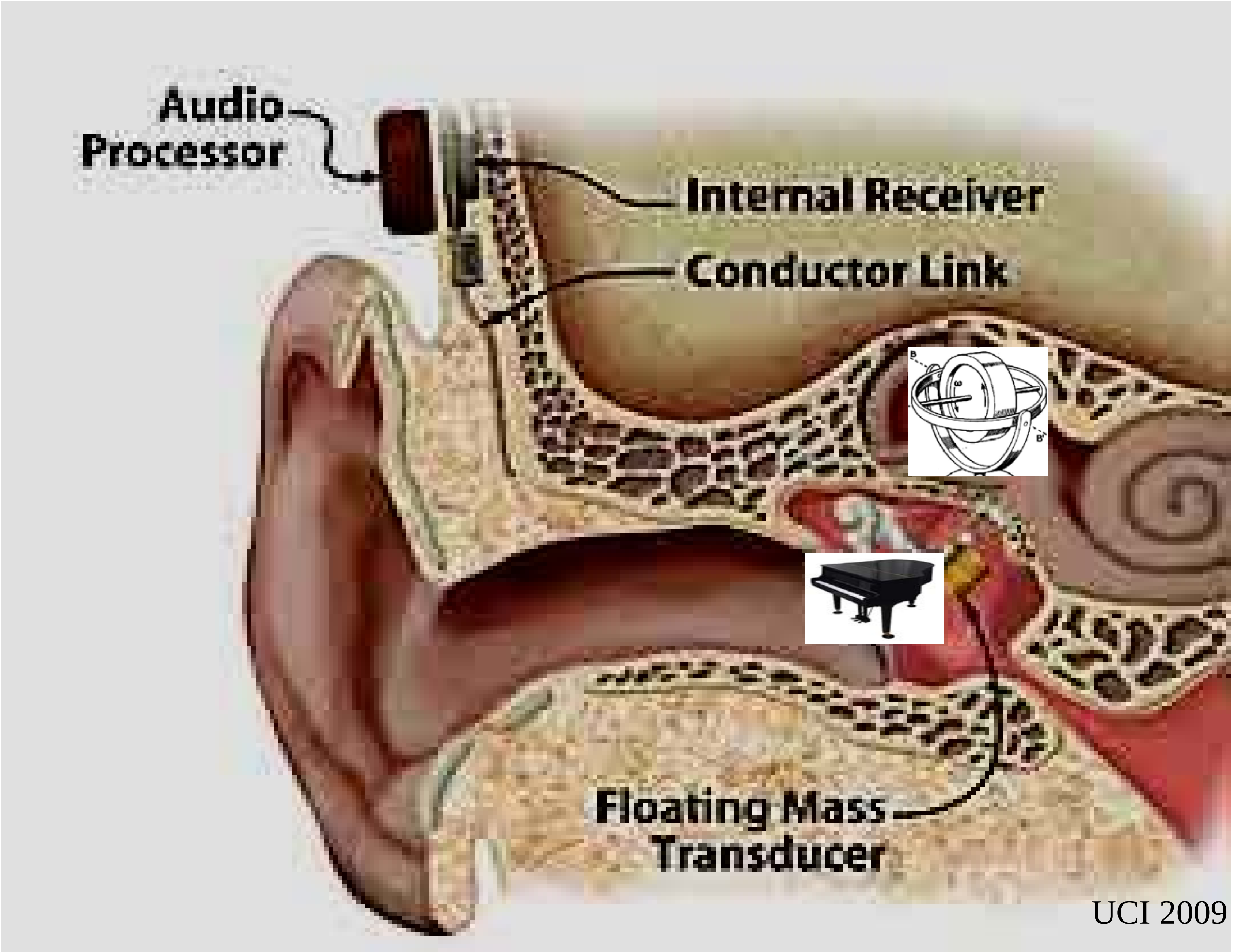
www.hearwireless.com

www.jabra.com

www.rnid.org.uk

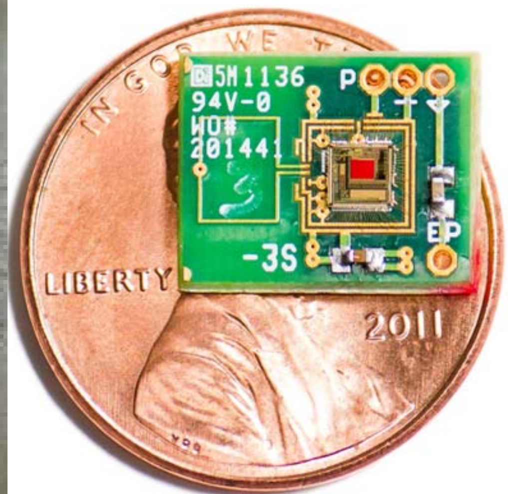
Turning iPhone into a hering aid

Micro- and Nano-Technology



UCI 2009

Inner ear powered radio

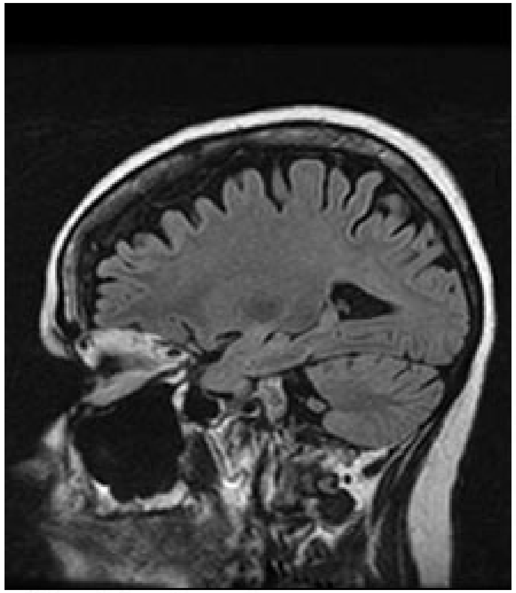


MIT Nature Nanotechnology 2012

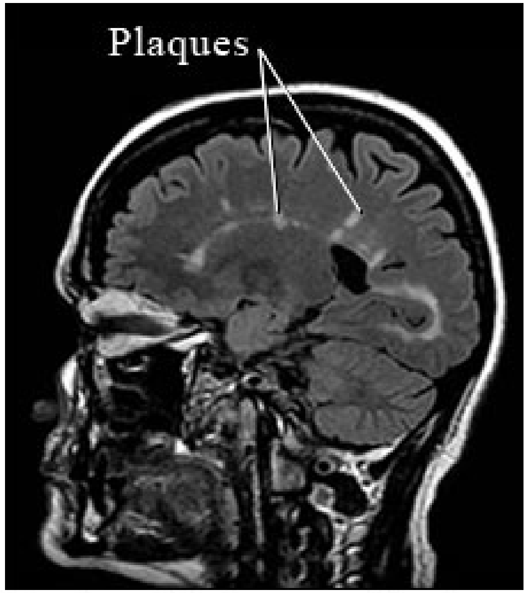
Applications for speech and hearing enhancement:

- Stuttering
- Aphasia
- Dyslexia
- Learning disability
- Multiple sclerosis
- Alzheimer’s disease
- Autism

…



Healthy brain



Brain with damage (lesions or plaques) caused by MS

Summary

- Chapter 1: Our work is interesting
- Chapter 2: Speech cues are redundant and complementary
- Chapter 3: Increasing functional signal-to-noise ratio is the key
- Chapter 4: Converging technologies to solve multiple problems