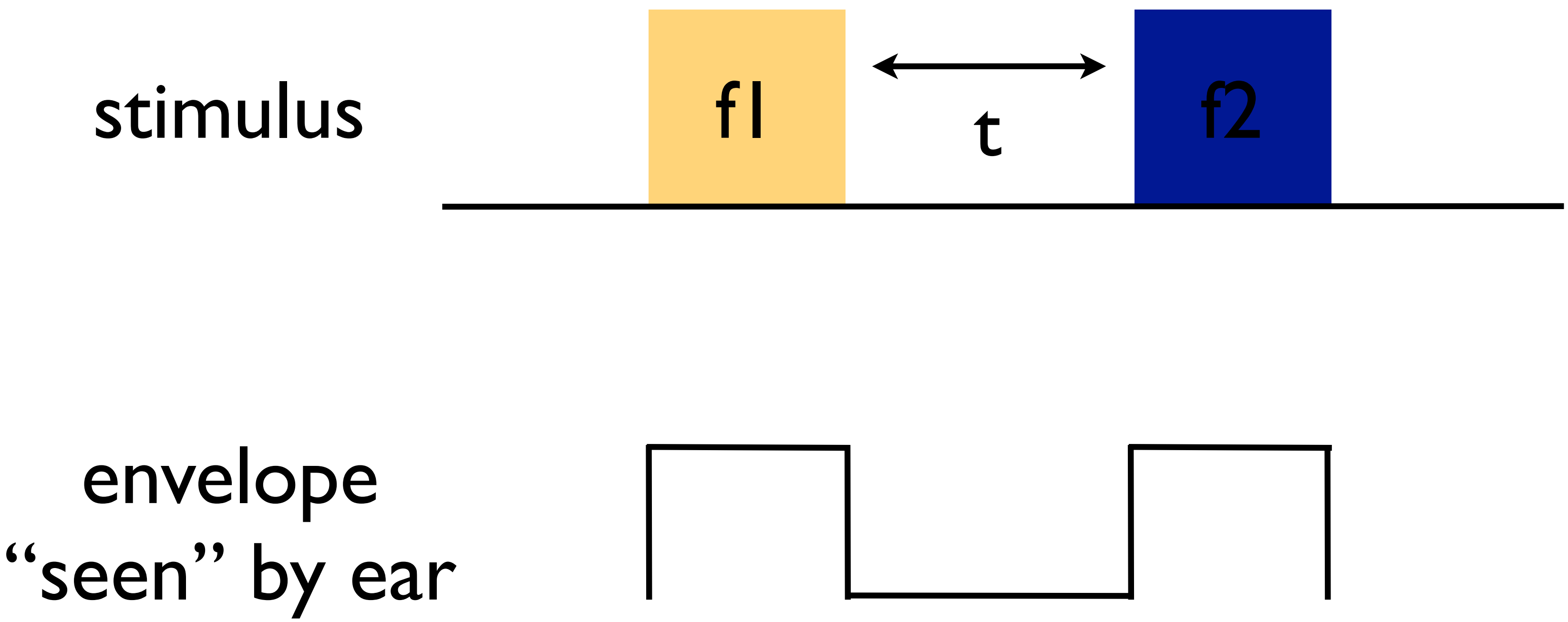
Modelling and signal processing of gap detection process

Willy Wong, University of Toronto

Acknowledgements

- *Speech ABR*
  Hilmi Dajani (assoc prof, U of Ottawa)
- *Method of significance*
  Dr. Jie Cui (Barrow Neurological Institute)
- *Neural modelling*
  Ewen Macdonald (assoc prof, TU Denmark)
  Sheena Luu

Gap detection



Gap in speech perception

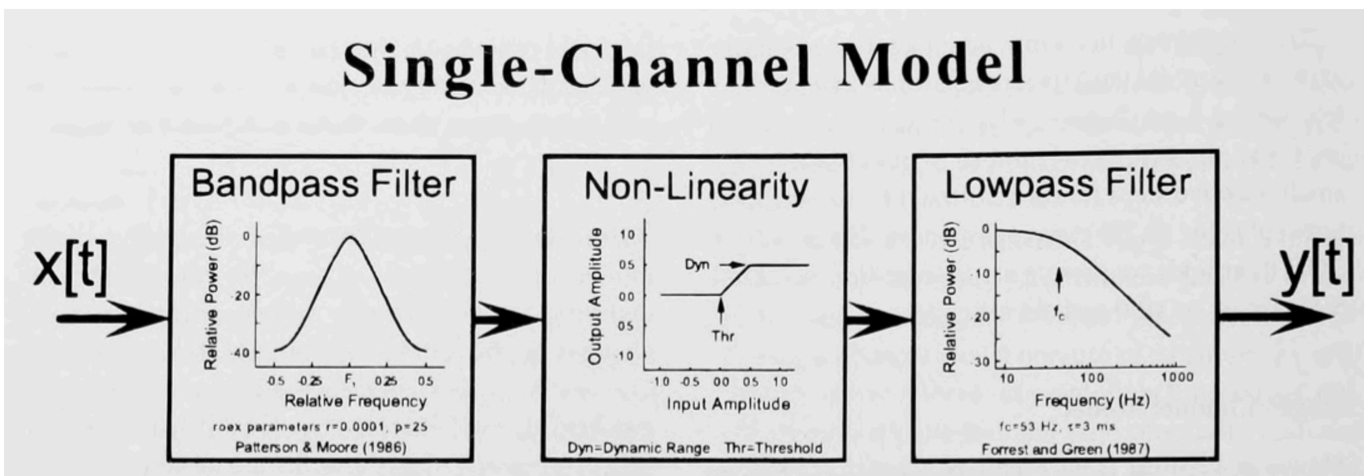
Gaps in speech signals can cue phonetically important distinctions; for example, a gap is one cue to the presence of the stop consonant phoneme /p/ in the word "spoon," in contrast to the word "soon" in which there is no gap.

Pichora-Fuller et al (2006)

Single-channel model

- Forrest et al (1991)
  Model of single-channel gap detection

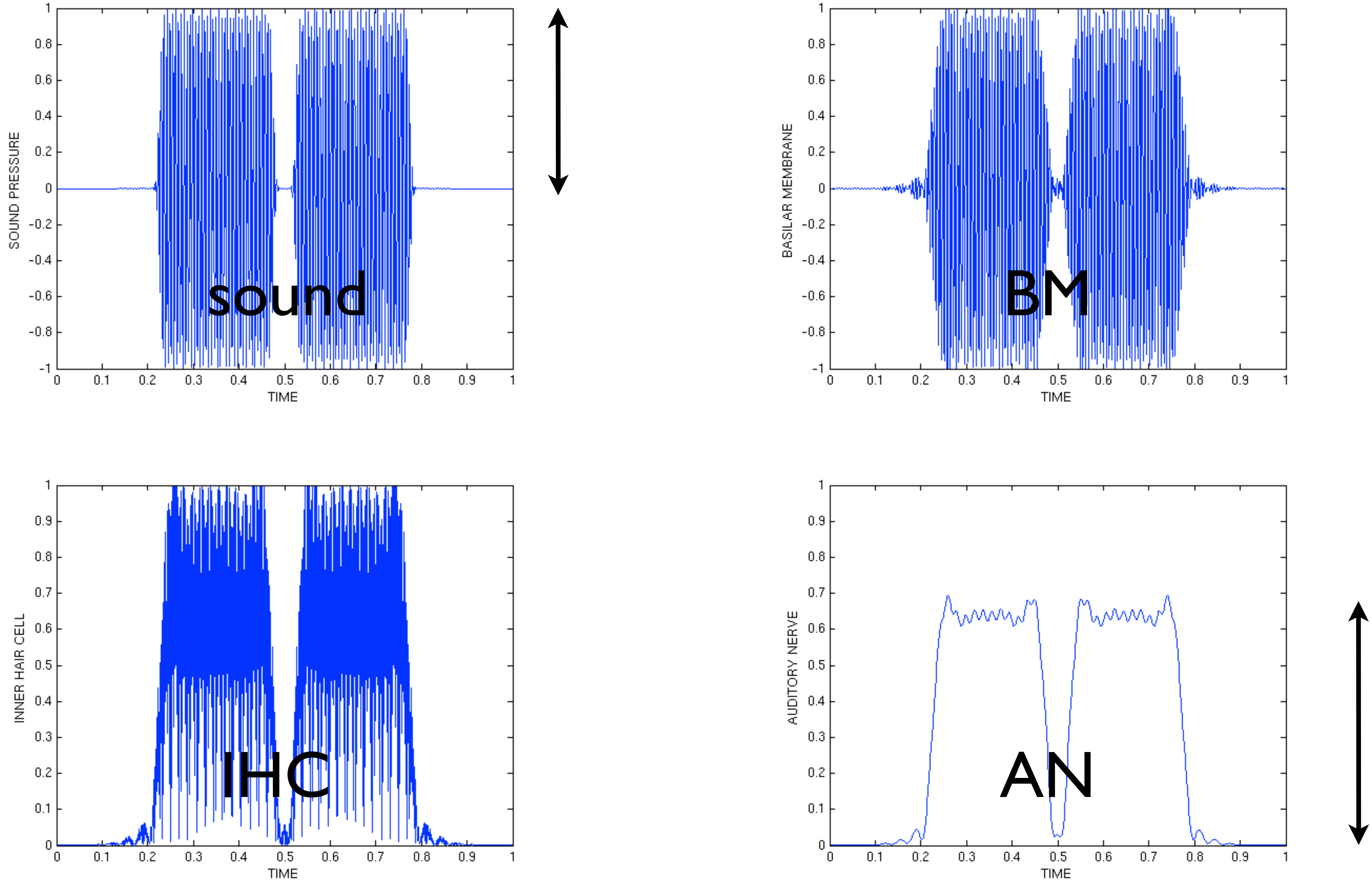
sound



towards central regions

BM IHC AN

Model response



Across channels

Formby et al (1998)
Across-channel gap detection

Small frequency: single-channel detection
Big frequency: between-channel detection

However, comprehensive model is missing

Electrophysiological approaches

- e.g. Shinohara et al (1995)
- correlation of ABR wave V to gap threshold
- EP threshold < PP threshold

We will be using an EP approach

Motivation

- Electrophysiology is informative but challenging to interpret/analyze
- Comprehensive model still missing
- *Hopefully there is a role for me here!*

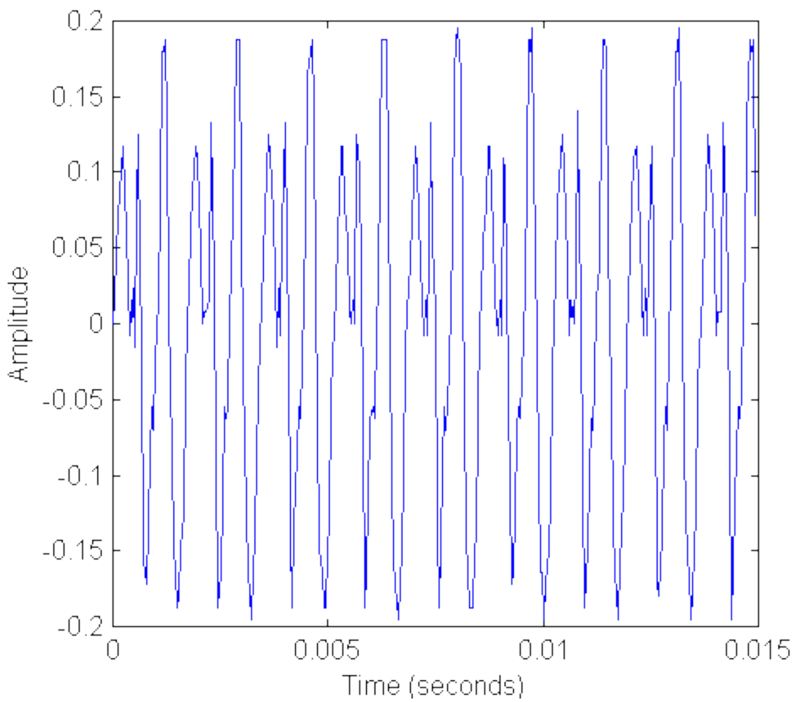
Outline

- Time-frequency analysis
- Blind statistical technique
- Theory and modelling

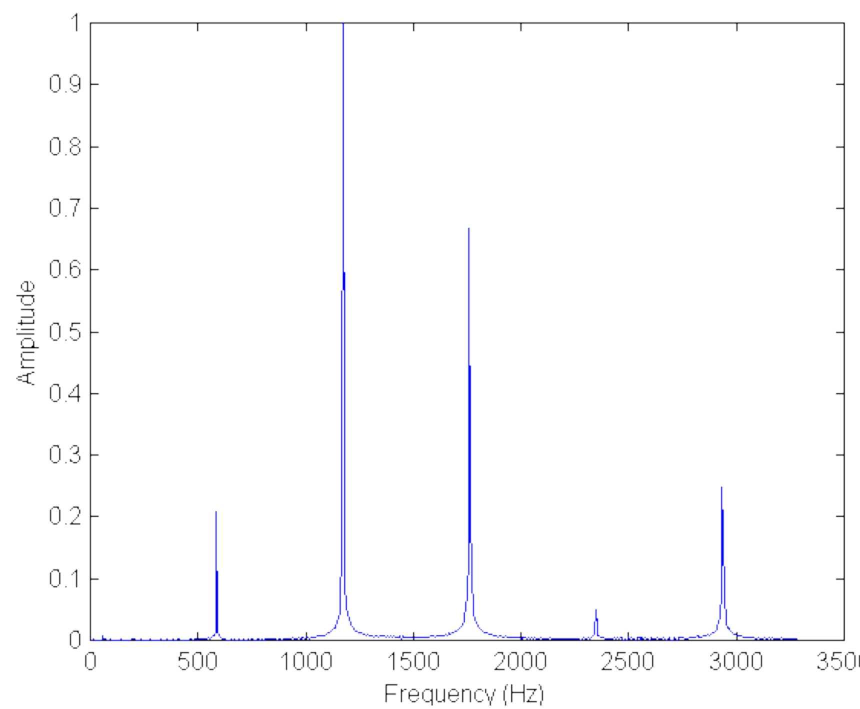
Time-frequency analysis

Frequency analysis and stationary signals

time domain

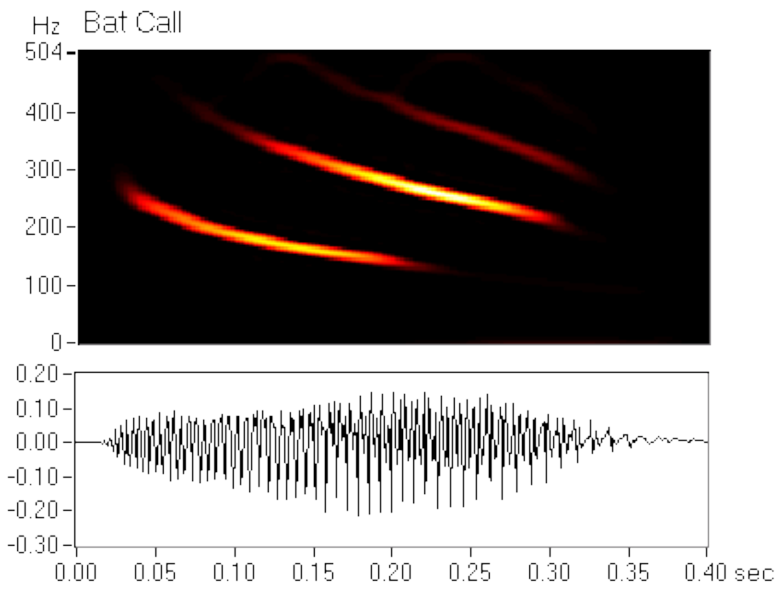


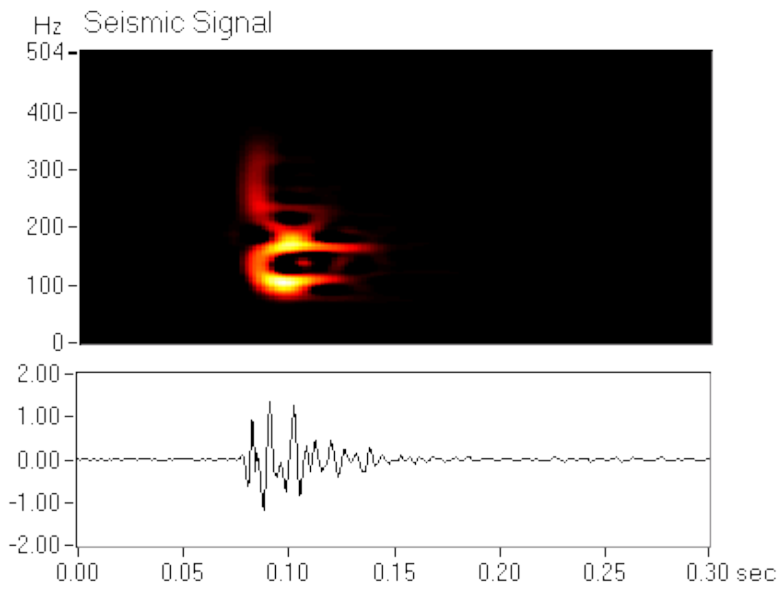
frequency domain



Non-stationary signals

Plotting spectral distribution over time is call a spectrogram

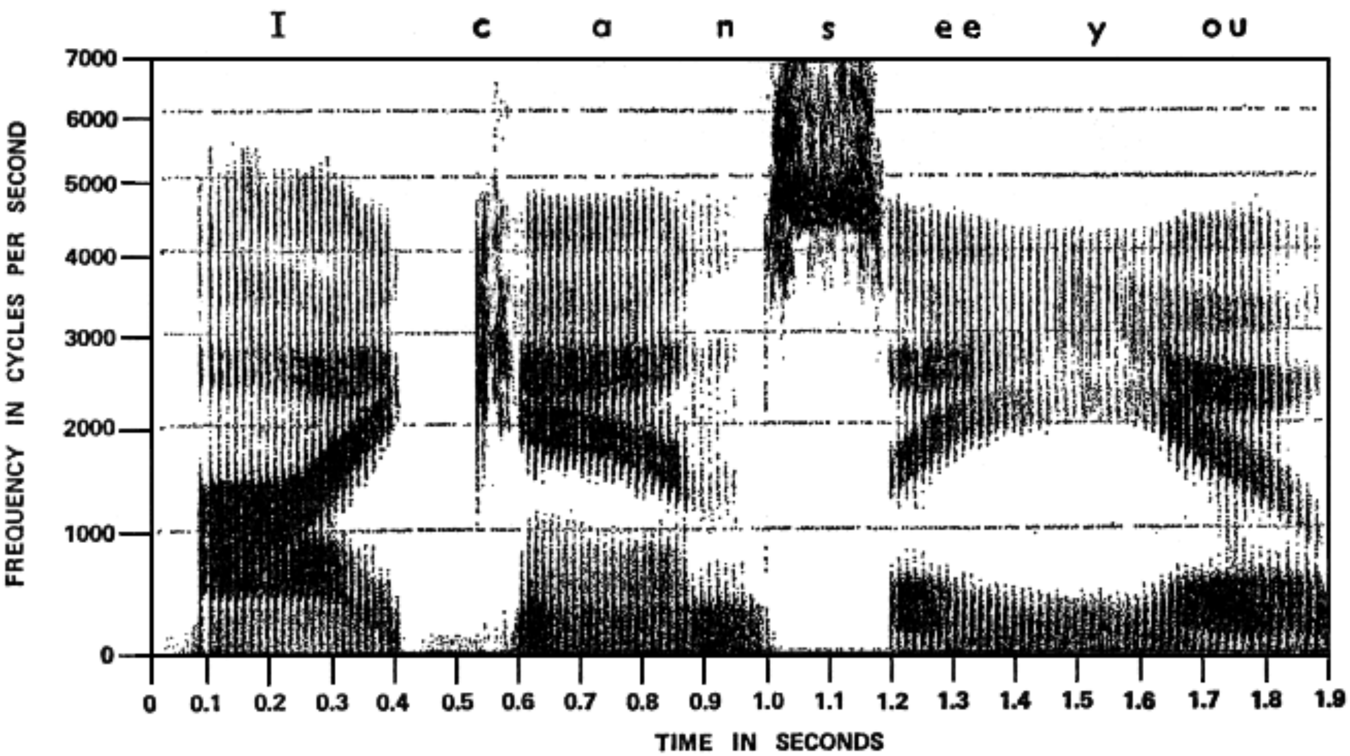




General area is called time-frequency analysis

Time-frequency analysis can applied to a variety of signals

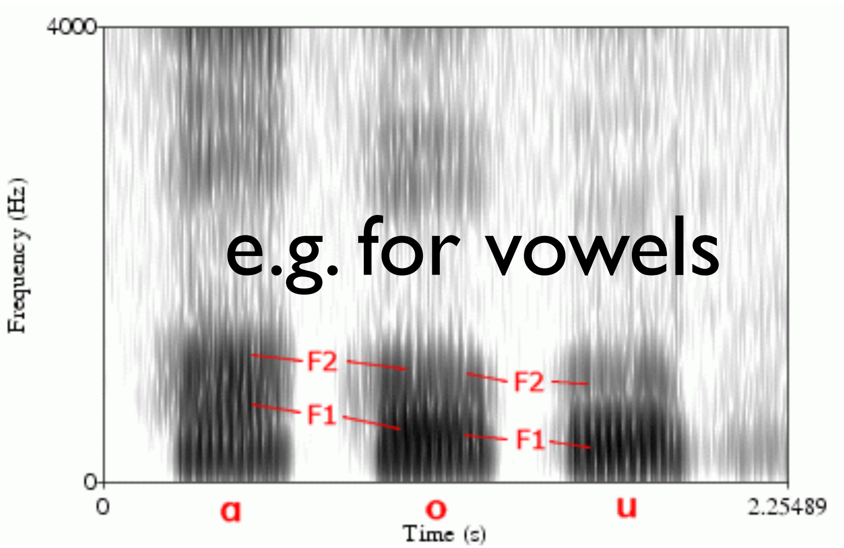
Speech spectrogram

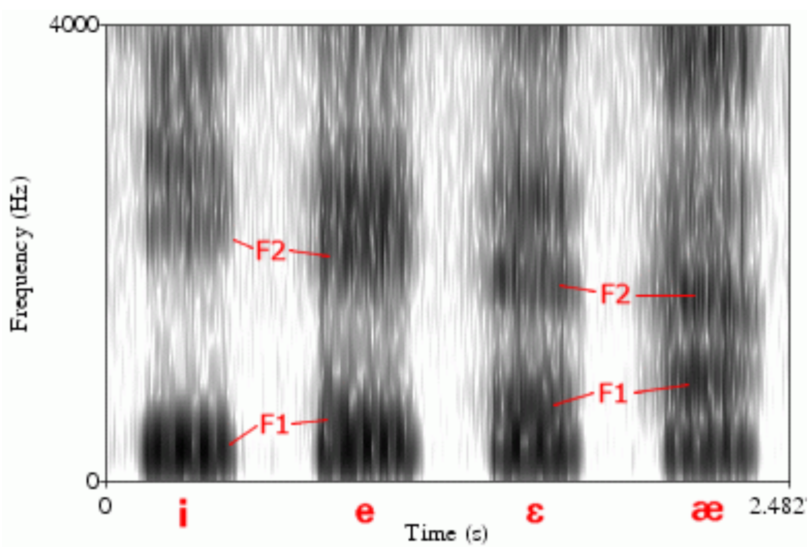


e.g. Winckel, *Music, Sound and Sensation*

Importance for formant analysis

e.g. for vowels

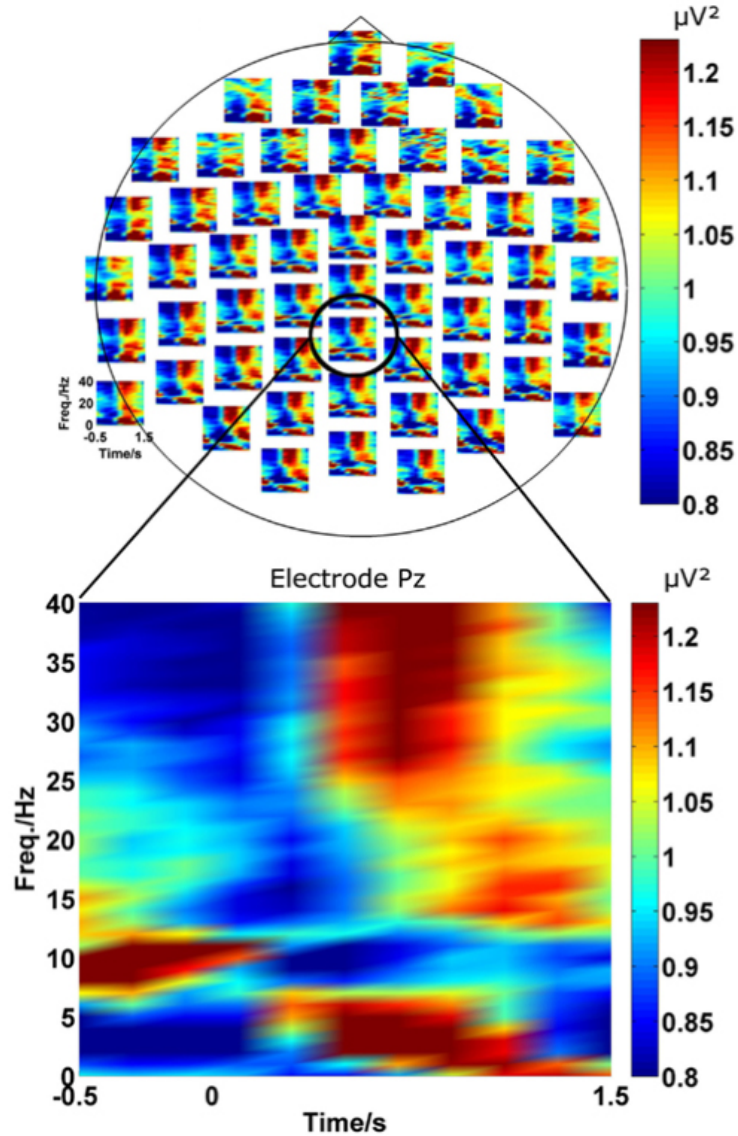




Average vowel formants[7]

Vowel (IPA)	Formant f_1	Formant f_2
i	240 Hz	2400 Hz
y	235 Hz	2100 Hz
e	390 Hz	2300 Hz
ø	370 Hz	1900 Hz
ɛ	610 Hz	1900 Hz
œ	585 Hz	1710 Hz
a	850 Hz	1610 Hz
æ	820 Hz	1530 Hz
ɑ	750 Hz	940 Hz
ɒ	700 Hz	760 Hz
ʌ	600 Hz	1170 Hz
ɔ	500 Hz	700 Hz
ɤ	460 Hz	1310 Hz
o	360 Hz	640 Hz
ɯ	300 Hz	1390 Hz
u	250 Hz	595 Hz

TF also good for EP data!

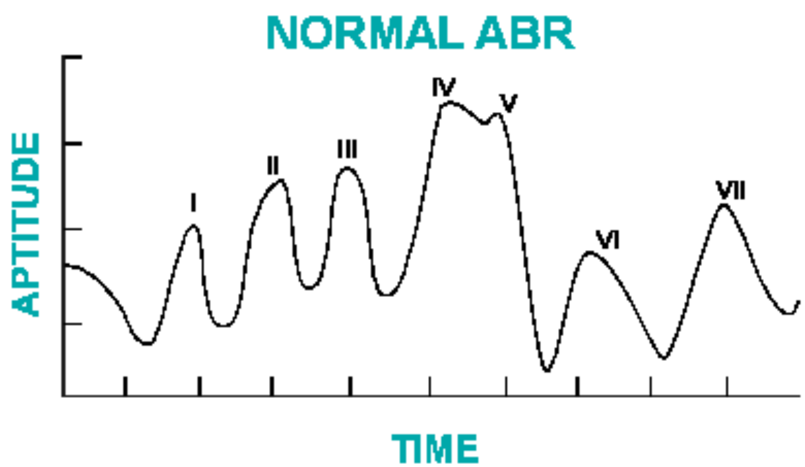


e.g. Farzan et al 2010

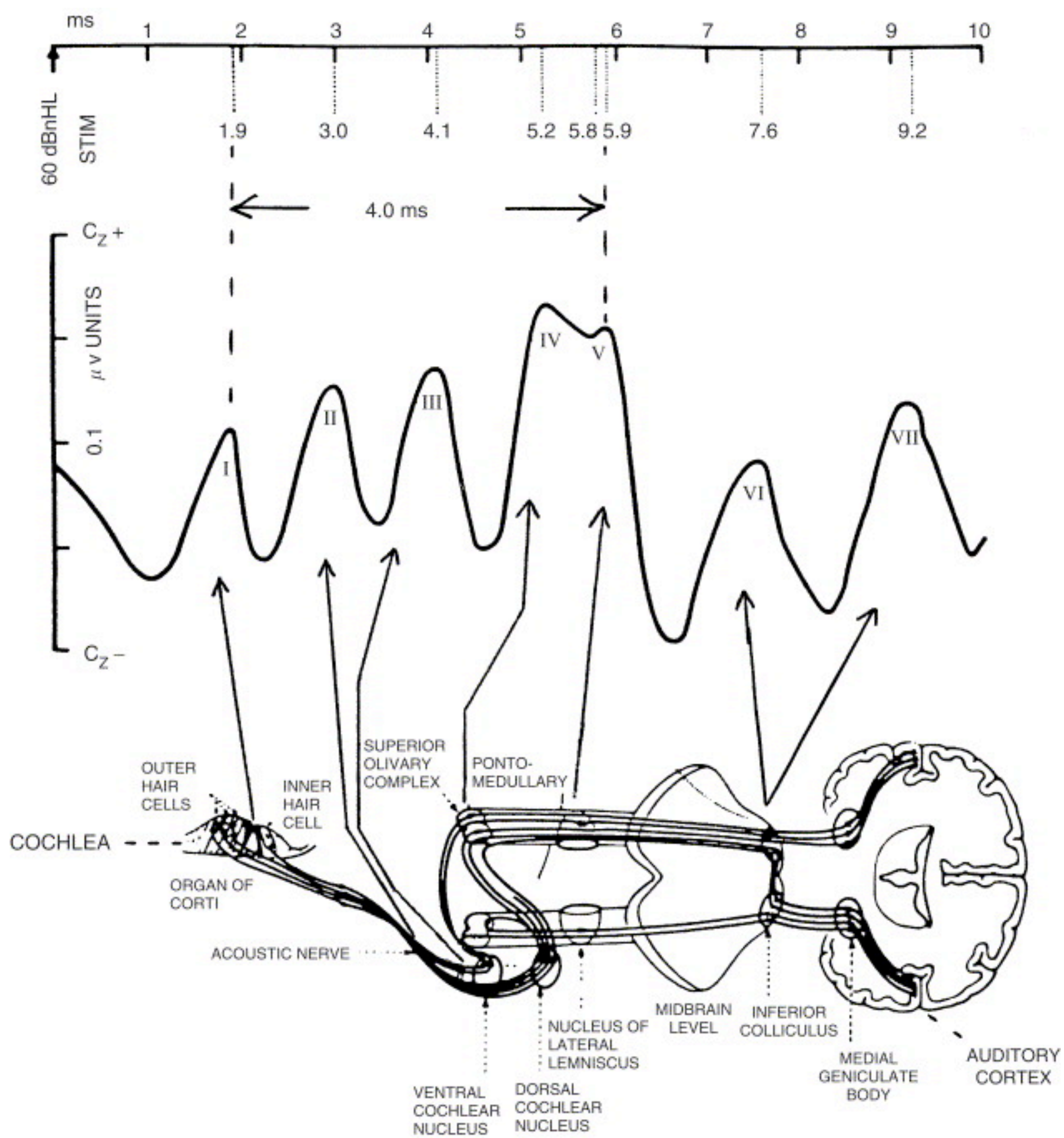
We did a relevant project involving time-frequency analysis of speech & ABR

Auditory brainstem response

"click" →



Periphery

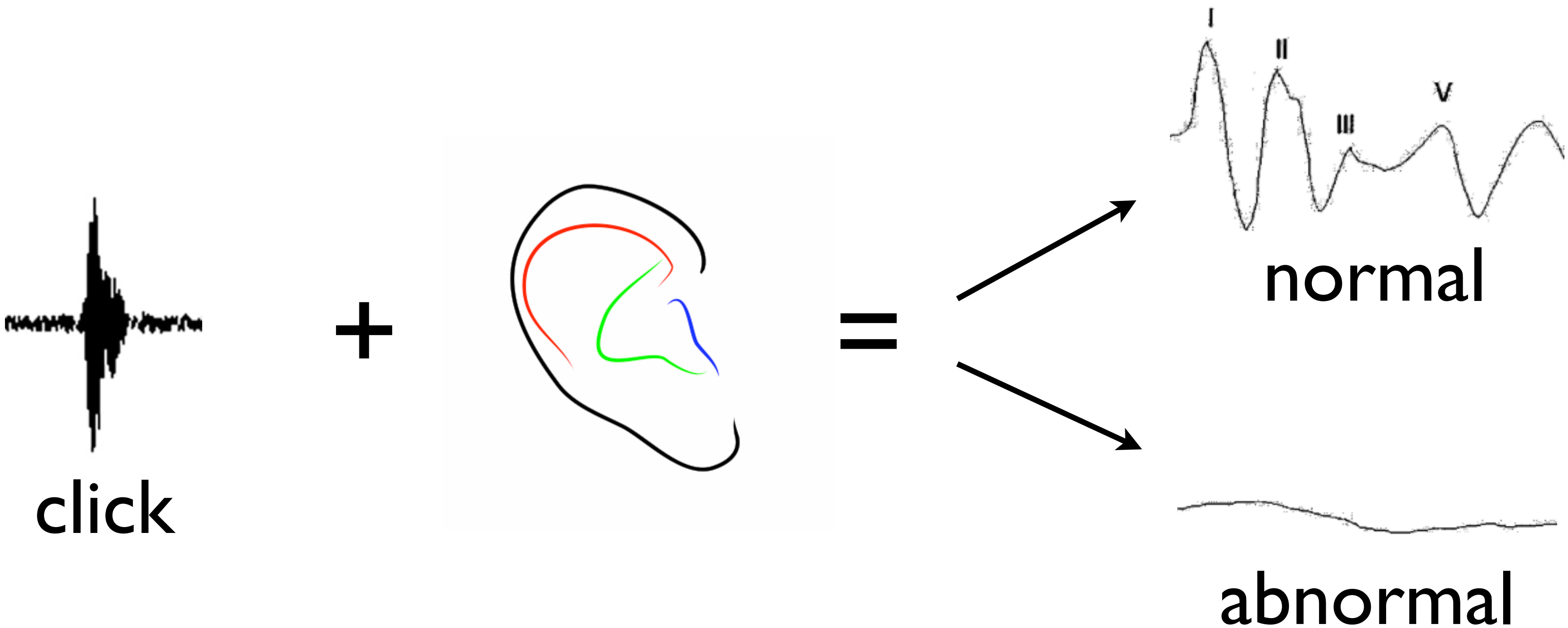


Brain

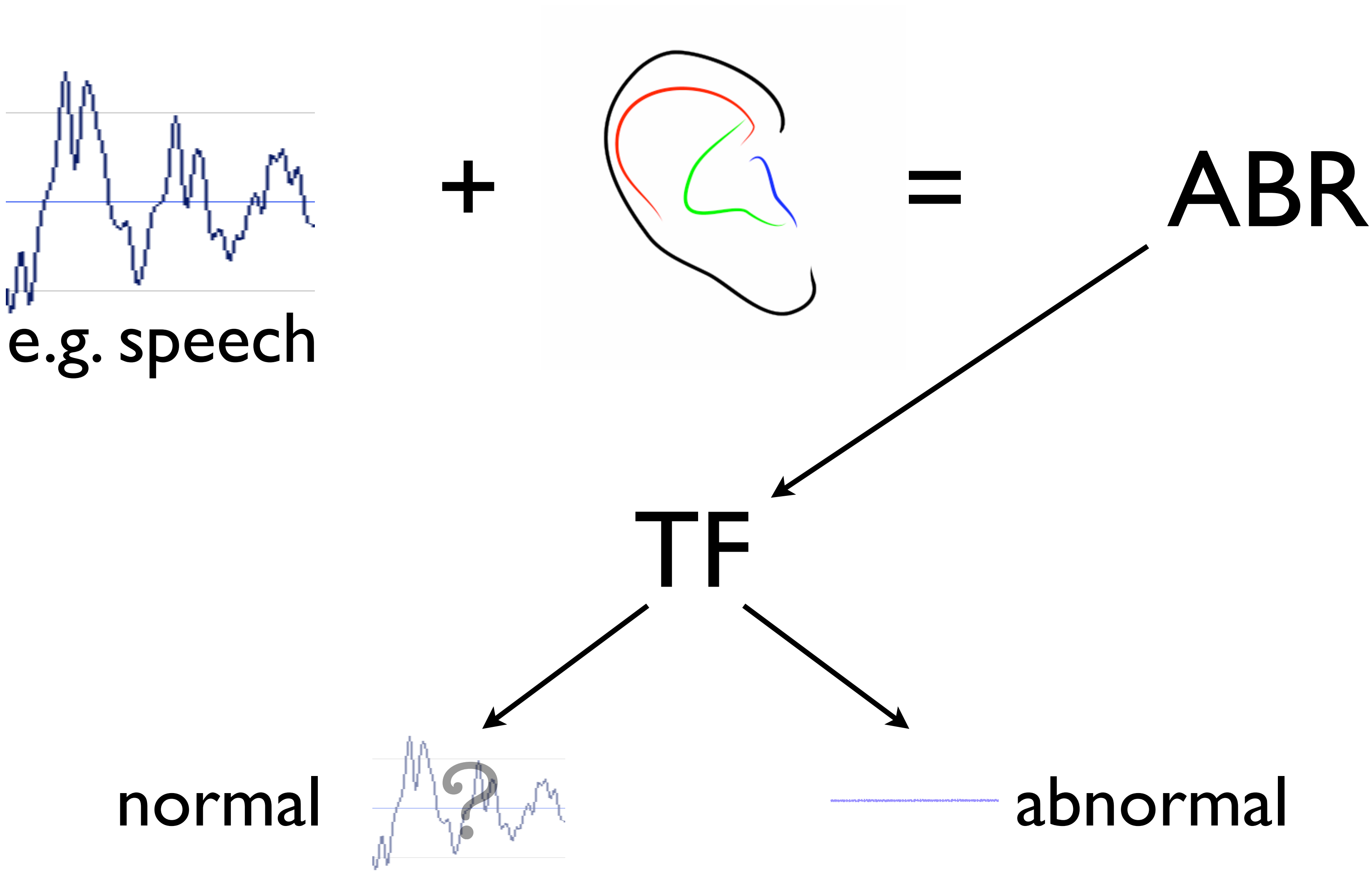
This is all fine if we are interested in "clicks".

But what about real signals like speech?

Clinical audiometry



Complex signal ABR



EP follows pitch contour

Recording Human Evoked Potentials That Follow the Pitch Contour of a Natural Vowel

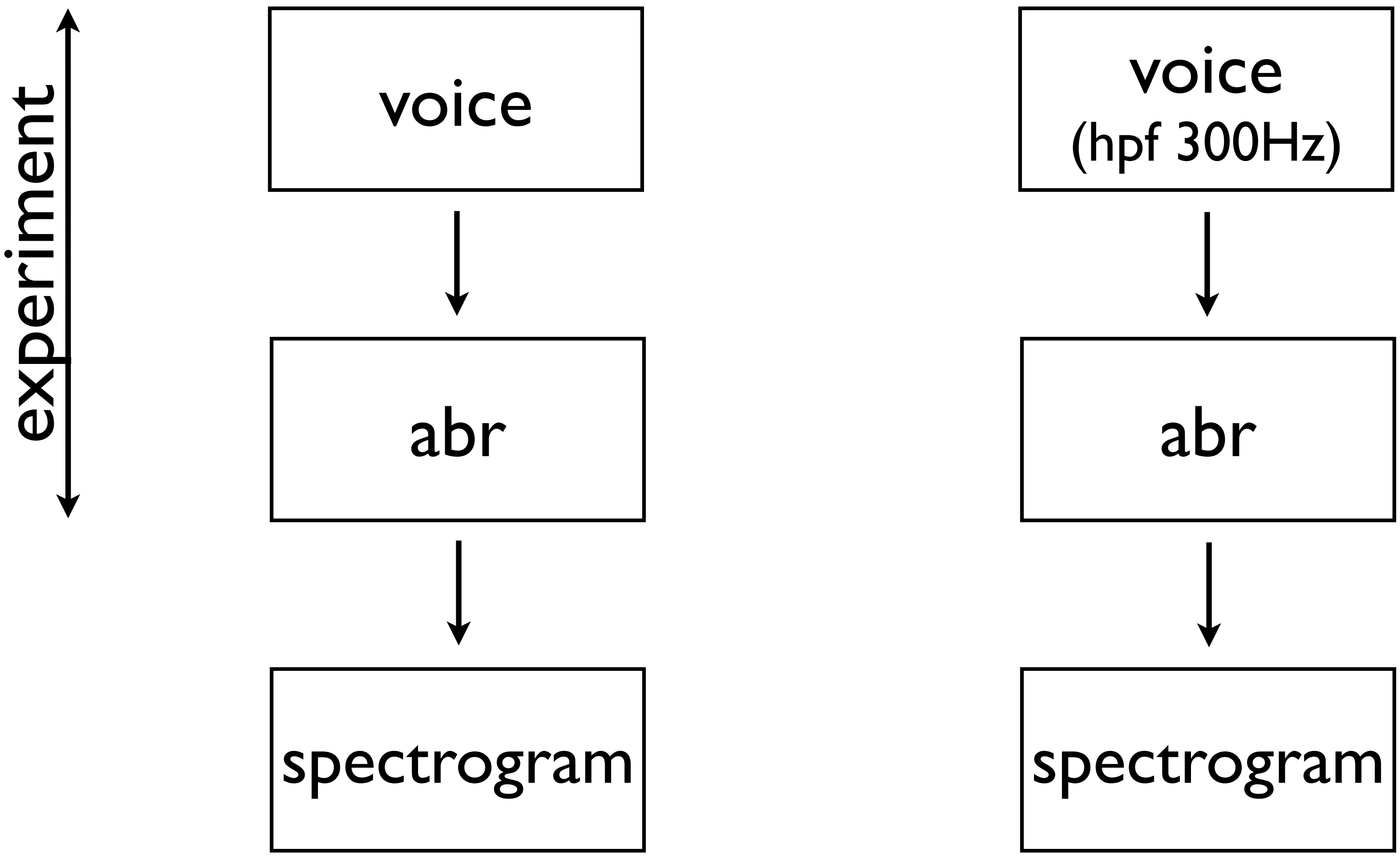
Hilmi R. Dajani*, David Purcell, Willy Wong, Hans Kunov, and Terence W. Picton

***Abstract*—We investigated whether pitch-synchronous neural activity could be recorded in humans, with a natural vowel and a vowel in which the fundamental frequency was suppressed. Small variations of speech periodicity were detected in the evoked responses using a fine structure spectrograph (FSS). A significant response ($P \ll 0.001$) was measured in all seven normal subjects even when the fundamental frequency was suppressed, and it very accurately tracked the acoustic pitch contour (normalized mean absolute error $< 0.57\%$). Small variations in speech periodicity, which humans can detect, are therefore available to the perceptual system as pitch-synchronous neural firing. These findings suggest that the measurement of pitch-evoked responses may be a viable tool for objective speech audiometry.**

***Index Terms*—Bioelectric potentials, biomedical signal processing, speech processing, time-frequency analysis.**

IEEE Transactions of Biomedical Engineering, 2005

Live demonstration

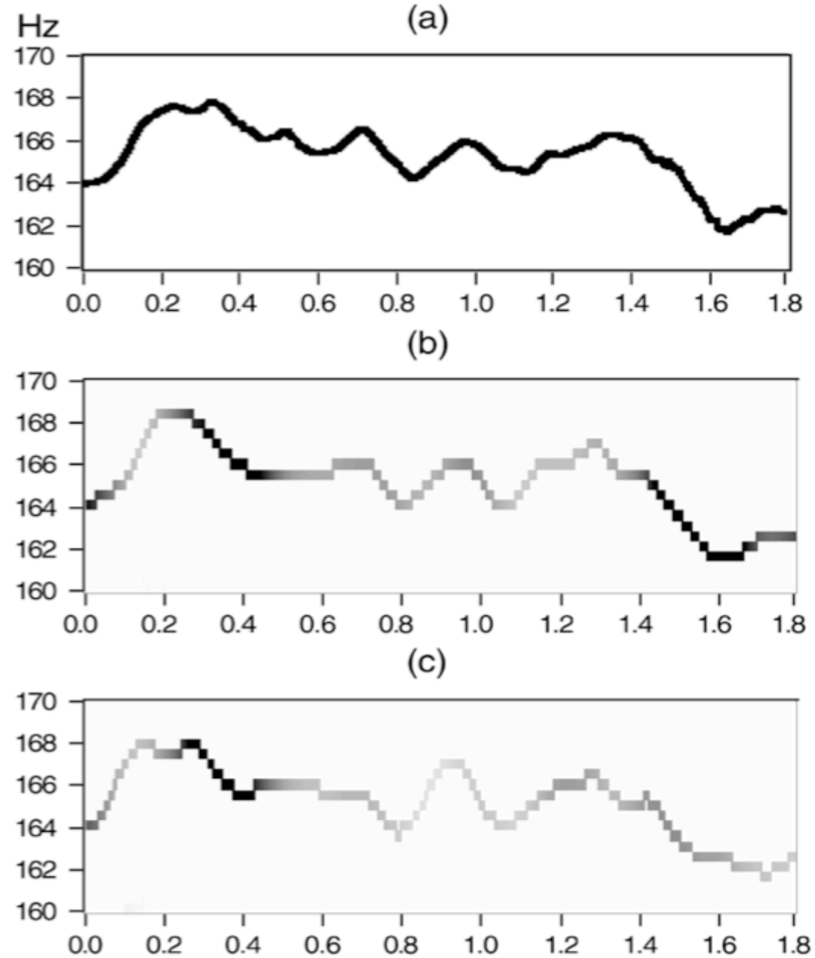


Speech ABR

Pitch track

Evoked response

Evoked response
(high-passed)



Repeated by other investigators e.g. cABR, Kraus (2010)

Research question

We can see complex stimulus in evoked response (in MEG as well?)

Similarly can we see "gap" in response?
If so, we can use to same method to study threshold!

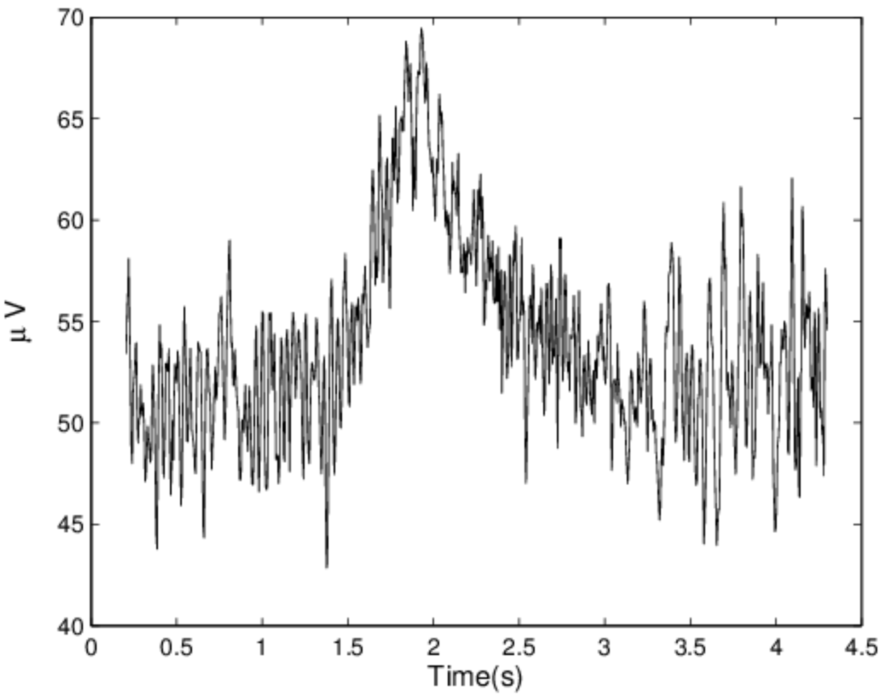
Method of significance

Motivation

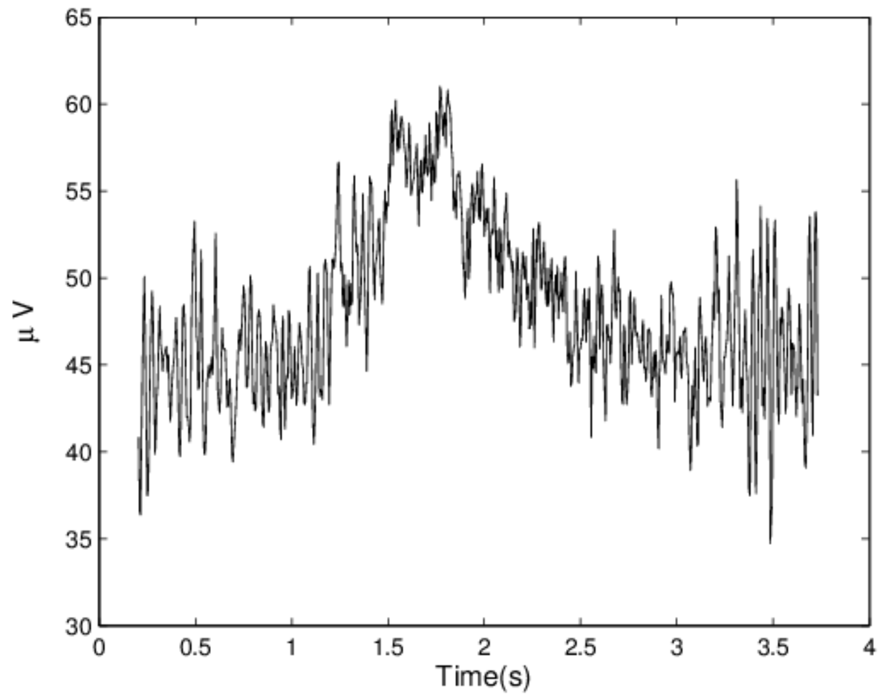
What if we don't know where to look?

Here is a "blind technique" that can help you

Neural signals are noisy

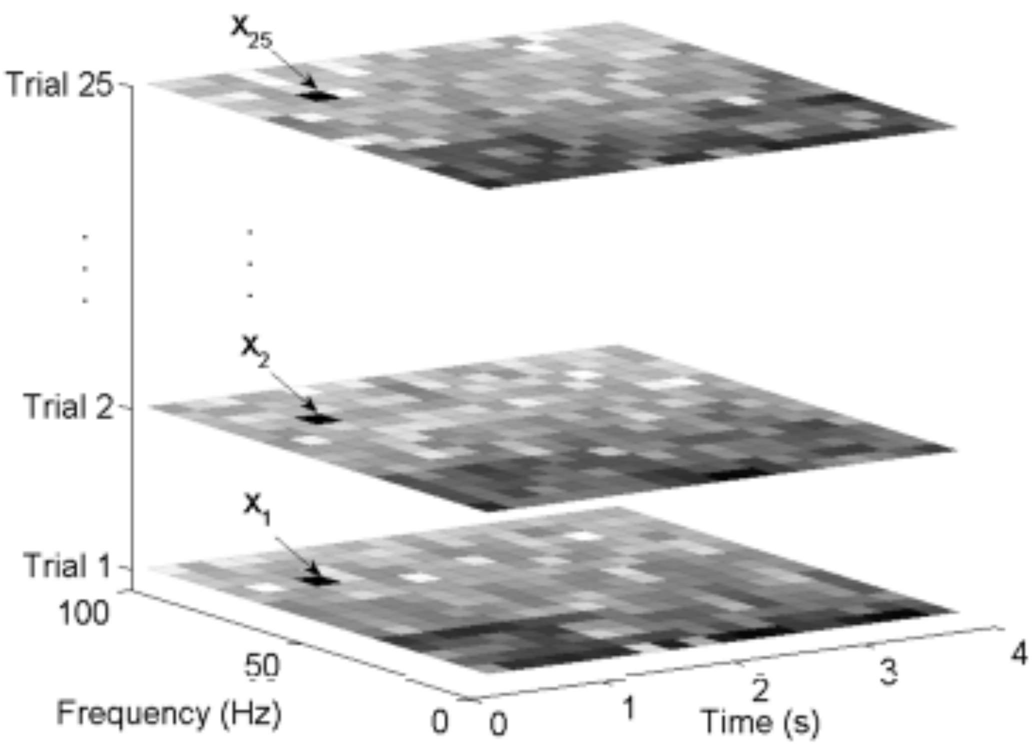


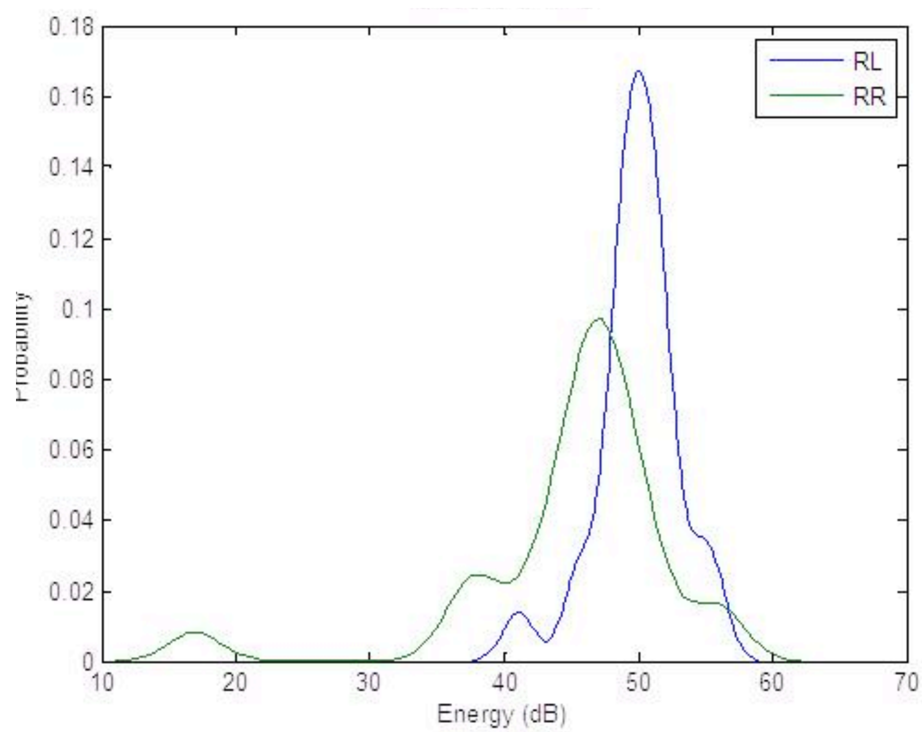
task A



task B

Method of significance





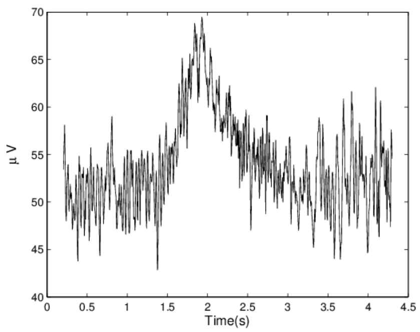
e.g. Talakoub et al 2013

Theory

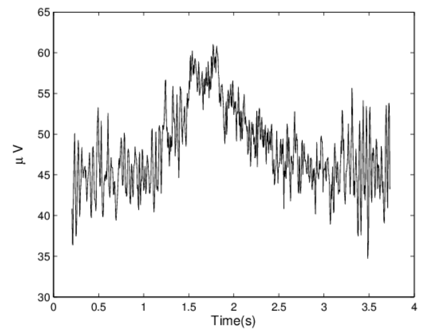
Typically x_i is power. $X : S_X = \{x_1, x_2, ..., x_N\}$
X is chi-squared distributed.

The transformation $Y = 10\log_{10}(X)$ gives a normal distribution.
T-test can be applied.

Test of significance can be used to find regions of difference in time-frequency space.

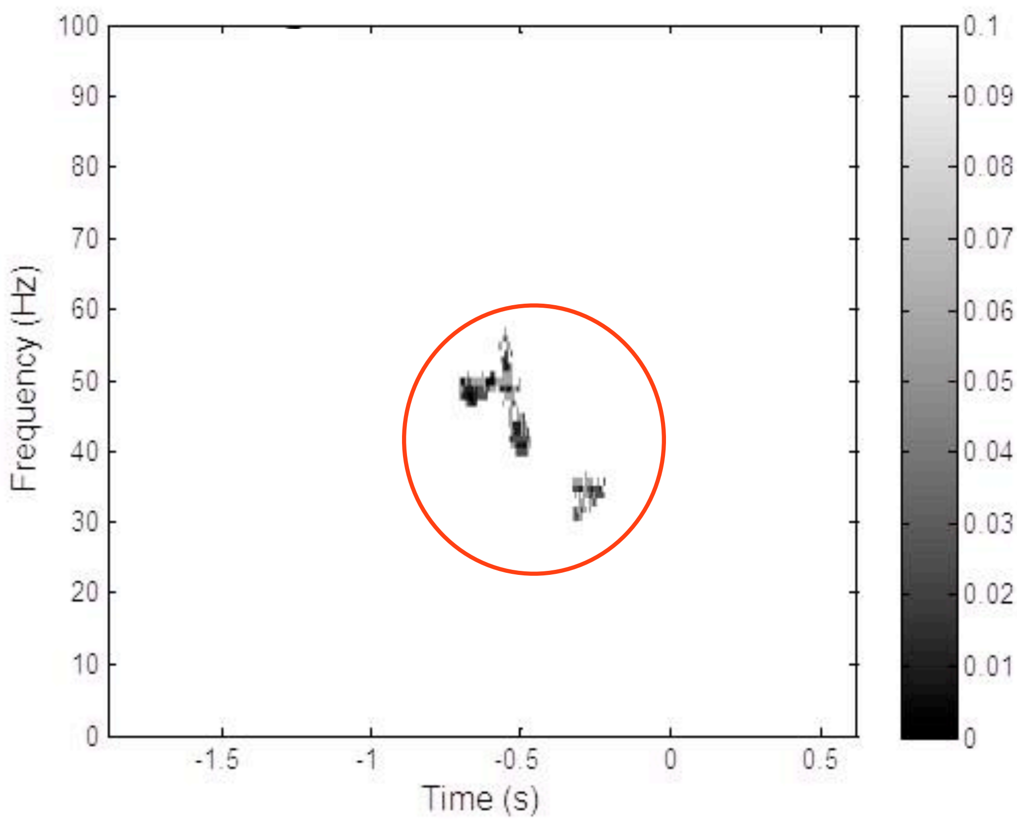


t−test



task A

task B



Research questions

Many investigators explore the P1-N1-P2 complex (50-200 ms after stimulus) for gap detection.

Could there be better correlates to threshold?

Method of significance can be used to find out!

Theory and modelling

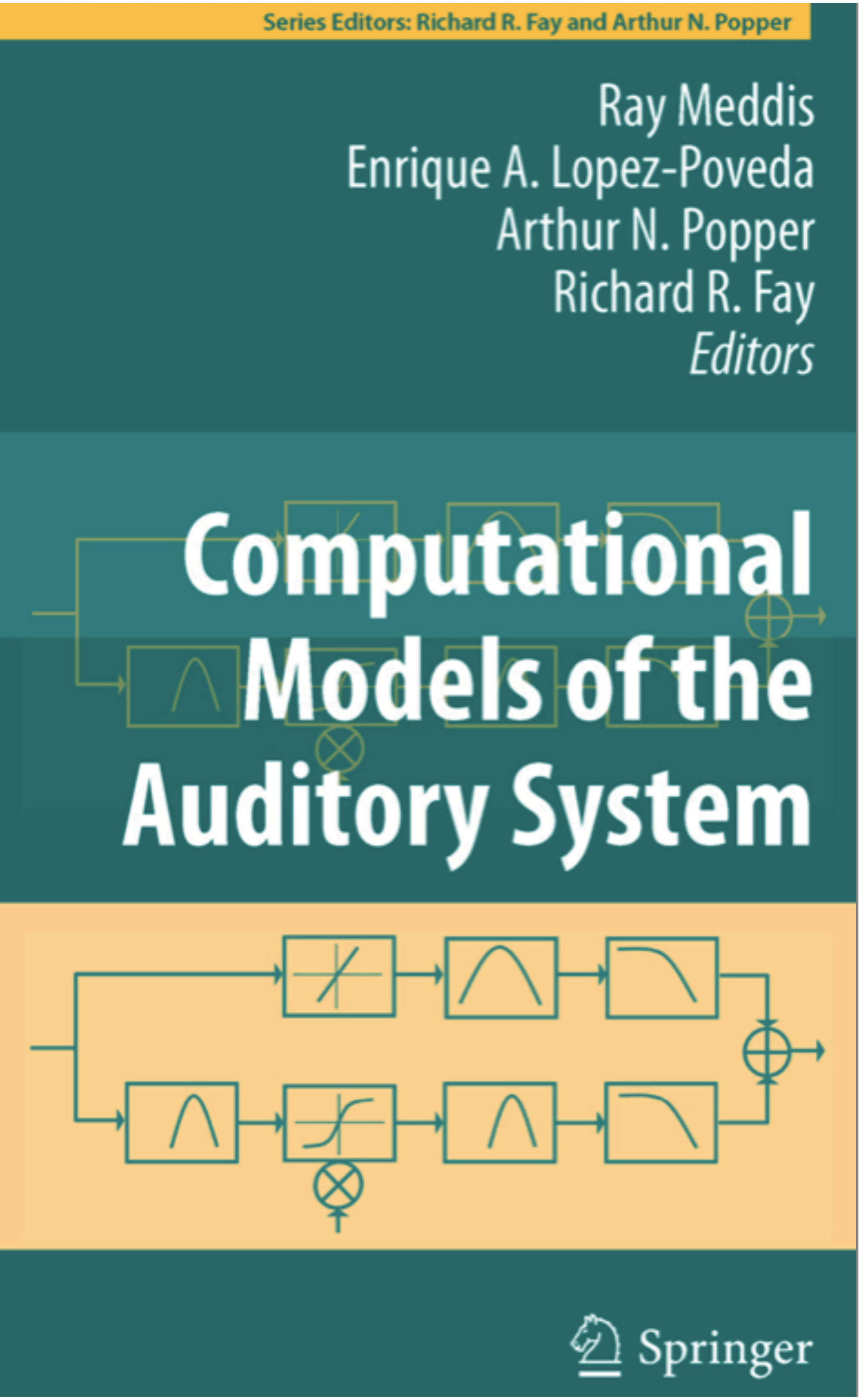
Aim

Can we have a comprehensive understanding of gap detection thresholds?

Review

- $F1 \simeq F2$: in-channel (peripheral processing)
- $F1 \neq F2$: between-channel (central processing)

***Need**: model of central processing of sound*

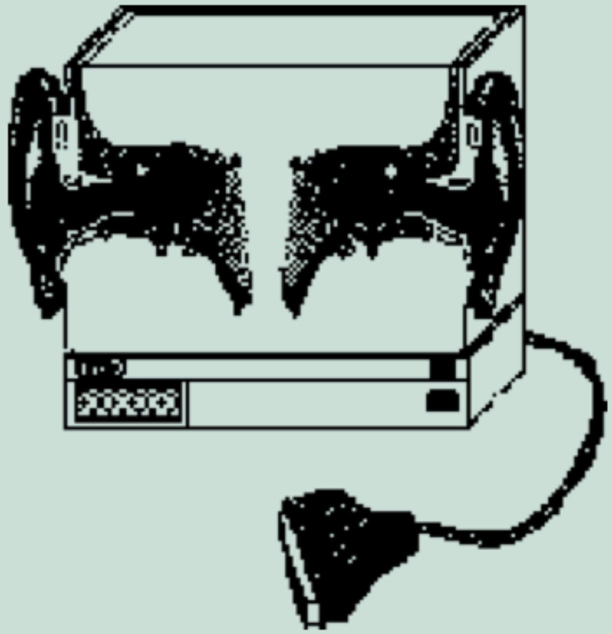

Series Editors: Richard R. Fay and Arthur N. Popper
Ray Meddis
Enrique A. Lopez-Poveda
Arthur N. Popper
Richard R. Fay
Editors
Computational Models of the Auditory System
Springer

Contents

DSAM

Development system for auditory modelling

Home Page



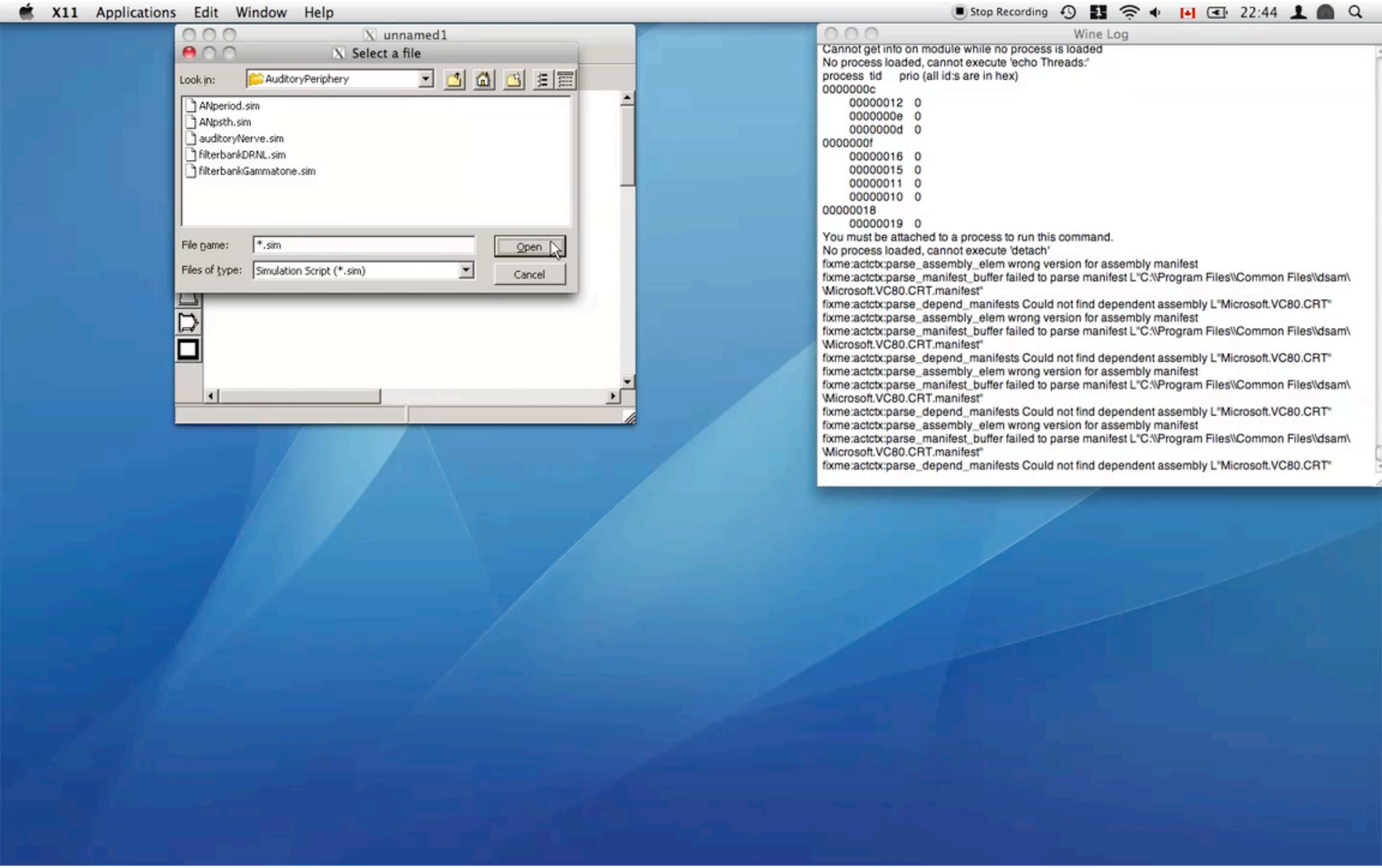
What is DSAM?
The Development System for Auditory Modelling (DSAM) is a computational library designed specifically for producing simulations ofthe auditory system. It brings together many established auditory models, produced by different research groups, under a flexible programming platform.

Why do you need DSAM?
It supports applications like AMS that run well known auditory models like the gammatone auditory filterbank, the Meddis hair cell, the Auditory Image Model (AIM) and others."

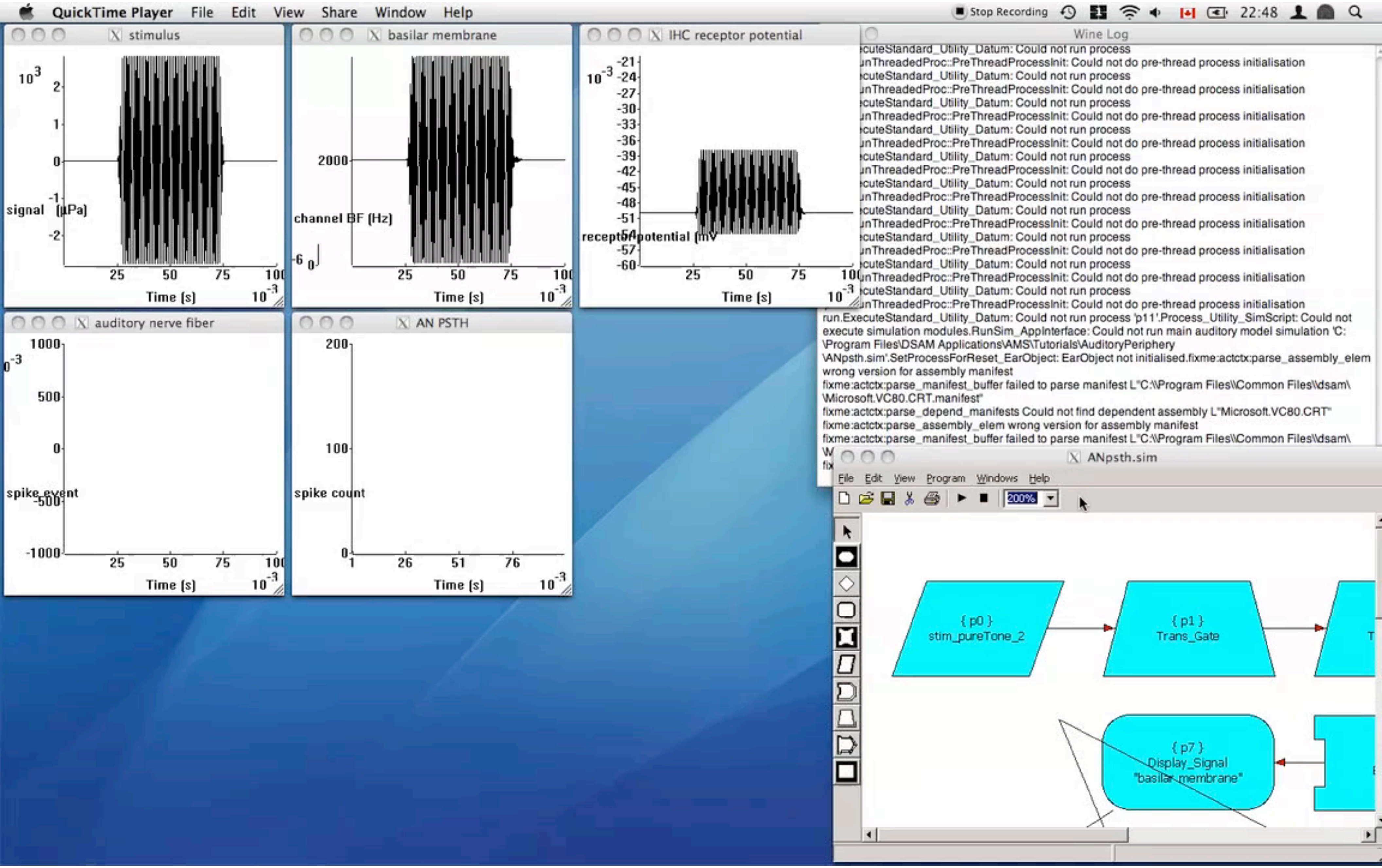
Site Contents
This site gives you access to the DSAM code, applications written using DSAM and links to related software.

- Introduction to DSAM
- DSAM Features
- The DSAM Policy
- Installing DSAM Applications
- Download Software
- Mailing list

DSAM is an open source project available from sourceforge. It was originally developed at the *Centre for the Neural Basis of Hearing (CNBH):*



X11
Applications
Edit
Window
Help
Stop Recording
22:44
unnamed1
Select a file
Look in:
AuditoryPeriphery
ANperiod.sim
ANpsth.sim
auditoryNerve.sim
filterbankDRNL.sim
filterbankGammatone.sim
File name:
*.sim
Open
Files of type:
Simulation Script (*.sim)
Cancel
Wine Log
Cannot get info on module while no process is loaded
No process loaded, cannot execute 'echo Threads:'
process tid prio (all id:s are in hex)
0000000c
00000012 0
0000000e 0
0000000d 0
0000000f
00000016 0
00000015 0
00000011 0
00000010 0
00000018
00000019 0
You must be attached to a process to run this command.
No process loaded, cannot execute 'detach'
fixme:actctx:parse_assembly_elem wrong version for assembly manifest
fixme:actctx:parse_manifest_buffer failed to parse manifest L"C:\\Program Files\\Common Files\\dsam\\Microsoft.VC80.CRT.manifest"
fixme:actctx:parse_depend_manifests Could not find dependent assembly L"Microsoft.VC80.CRT"
fixme:actctx:parse_assembly_elem wrong version for assembly manifest
fixme:actctx:parse_manifest_buffer failed to parse manifest L"C:\\Program Files\\Common Files\\dsam\\Microsoft.VC80.CRT.manifest"
fixme:actctx:parse_depend_manifests Could not find dependent assembly L"Microsoft.VC80.CRT"
fixme:actctx:parse_assembly_elem wrong version for assembly manifest
fixme:actctx:parse_manifest_buffer failed to parse manifest L"C:\\Program Files\\Common Files\\dsam\\Microsoft.VC80.CRT.manifest"
fixme:actctx:parse_depend_manifests Could not find dependent assembly L"Microsoft.VC80.CRT"
fixme:actctx:parse_assembly_elem wrong version for assembly manifest
fixme:actctx:parse_manifest_buffer failed to parse manifest L"C:\\Program Files\\Common Files\\dsam\\Microsoft.VC80.CRT.manifest"
fixme:actctx:parse_depend_manifests Could not find dependent assembly L"Microsoft.VC80.CRT"


QuickTime Player File Edit View Share Window Help
Stop Recording
22:48
stimulus
signal (µPa)
Time (s)
basilar membrane
channel BF (Hz)
Time (s)
IHC receptor potential
receptor potential (mV
Time (s)
Wine Log
auditory nerve fiber
spike event
Time (s)
AN PSTH
spike count
Time (s)
ANpsth.sim
File Edit View Program Windows Help
200%
{ p0 }
stim_pureTone_2
{ p1 }
Trans_Gate
{ p7 }
Display_Signal
"basilar membrane"

Research questions

- Simulate *peripheral neural response* for gap detection
- Extend method to *between-channel gap detection in other domains (speech, localization, etc)*
- Can study relative timing between channels

Conclusions

We hope some of these tools will be useful for project

Happy to collaborate on data analysis and model development!

Questions?