



Phase ~~and Feedback~~ in the Nonlinear Brain

Malcolm Slaney (IBM and Stanford)

Hiroko Shiraiwa-Terasawa (Stanford)

Regaip Sen (Stanford)

Auditory processing pre-COSYNE workshop
March 23, 2004



Simplistic Models

- “Dehydrated cats and the application of Fourier analysis to hearing problems become more and more a handicap for research in hearing.”—von Bekesy
- Overly simplistic models are bad
 - Wrong path
 - For example, critical band filters
 - Pitch and phase



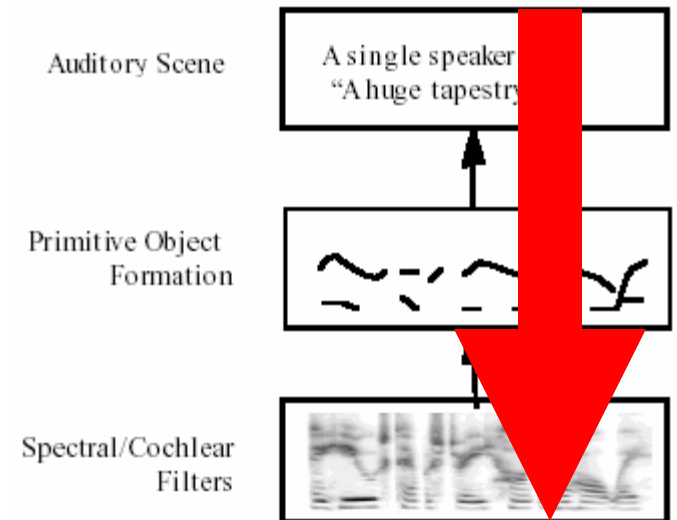
Tony's Goals



- Linearity
 - Taylor says everything is linear
 - To a point
 - Temporal codes
 - I don't know the answer....
 - Attentional effects
 - Big problem.. Not today
 - Phase perception

A Critique of Pure Audition

- Marr
 - Pure vision
- Churchland
 - A critique of pure vision
- Slaney
 - A critique of pure audition
 - Where are efferents?



Motivation

Speech processing in the auditory system: The representation of speech sounds in the responses of the auditory nerve

Shihab A. Shamma¹
 Department of Electrical Engineering, University of Maryland, College Park, Maryland 20742
 (Received 3 January 1985; accepted for publication 8 July 1985)

In a previous paper the speech evoked spatio-temporal response patterns recorded in large populations of auditory-nerve fibers in the cat were examined (M. J. Miller and M. B. Sachs, *J. Acoust. Soc. Am.* 74, 202-517 (1983)). The distribution of the phase of synchronized activity emerges as an important response feature reflecting the stimulus spectral parameters. Specifically, each strong low-order harmonic of the stimulus (< 1.5-2 kHz) domain synchrony of a relatively broad segment of fibers near its corresponding cochlear nucleus (CN) location in a pattern which mirrors the underlying traveling wave component fiber segment can be roughly subdivided into two regions: (1) a region based to the resonance of the harmonic where the fiber PST histograms accumulate only small shifts relative to each other reflecting the fast speed of propagation of the traveling wave region at or very near the point of resonance where the responses exhibit drastic shifts owing to the sudden slow down of the traveling wave and the consequent rapid accumulation of phase shifts. These rapid phase shifts that manifest themselves as localized spatial discontinuities in an otherwise relatively uniform instantaneous phase activity across the fiber array, all occurring at the CN locations corresponding to the harmonics of the stimulus.

PACS numbers: 43.63.Pd

INTRODUCTION

The sounds of speech evoke complex patterns of activity in the peripheral auditory system. This activity codes in a way meaningful to the central nervous system (CNS) the perceptually significant parameters of the signal. The multidimensionality of these percepts must, in general, involve a large number of such parameters. However, for the perception of some voiced speech utterances, e.g., vowel-like sounds and voiced consonant-vowel combinations, a minimal sufficient set of descriptors has been established (Funn, 1971); they include the formant frequencies and their transitions, and the low-frequency harmonic components which are important for the perception of pitch (Ritsma, 1968). The coding of these signal parameters in the responses of the auditory nerve has been a central theme in the neurophysiology of the peripheral auditory system (DeGruete and Kiang, 1984; Evans, 1978; Kiang, 1980; Sachs and Young, 1979; Sines and Oertler, 1983). Two questions have been posed: What are the response features that code these parameters? How is the central nervous system able to exploit them and hence, recognize the underlying stimulus? Recent studies have gone a long way toward answering the first question. Several response measures and properties have emerged as potential cues from which the CNS may derive the appropriate percept, among them the temporal periodicities, average firing rates, and their distributions across the auditory-nerve fiber array (DeGruete, 1984; Young and Sachs, 1979). In this report we examine the expression of such cues in the spatial and temporal profiles of auditory-nerve responses. In the companion paper (Shamma, 1985) we

¹Part of this work was conducted while the author was a staff fellow at the Mathematical Research Branch, NIDH, and earlier in his Section on Comparative Biology, NIH.

address the question of auditory stimulus spectral parameters. The complex patterns of activity have two important aspects: excite fibers innervating distototopically ordered spatial spans of fibers to low-frequency bands looked to the wave. Fig. 1 will illustrate schematically the auditory-nerve response to a stimulus (e.g., the vowel /a/ low-order harmonics of the along the cochlea. The stimulus is those located near the formant frequencies of the fiber. The fiber at CF = 0.41, 0.62 harmonic, first formant, and Phase locking diminishes as Phase locking diminishes as may still lock to the envelope of the interfering harmonics which the fiber at CF = 2.22 kHz patterns of the tonotopical array are then projected to networks [e.g., Lateral Inhibitory Networks (LIN)] prior to further analysis. Experimental data have recently been available that allow a reasonable reconstruction of such spatio-temporal response patterns from single unit recordings of large populations of auditory-nerve fibers (Sachs and Young, 1979; Young and Sachs, 1979; Miller and Sachs, 1983; Burta, 1985). These plots summarize coarsely the nerve responses and facilitate in particular the examination of the spatial features that result from phase changes in its synchronous responses. In the following sections, we will examine these features in terms of the evoked by these complex

S. Shamma. "Speech processing in the auditory system." 1985.

Good Shihab

An account of monaural phase sensitivity

Robert P. Carlyon¹
 MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom

Shihab Shamma
 Electrical and Computer Engineering, Institute for Systems Research, University of Maryland, College Park, Maryland 20742

(Received 3 January 2002; revised 30 March 2003; accepted 7 April 2003)

Listeners can detect phase differences between the envelopes of sounds occupying remote frequency regions, and between the fine structures of partials that interact within a single auditory filter. They are insensitive to phase differences between partials that occur for all three of these findings.

A new model is proposed that can account for all three of these findings. A widely popular approach, does not discuss across-channel timing differences. It is predicted quantitatively by analyzing the output of a cochlear model composition inspired by responses of neurons in the auditory cortex, and metrics between the responses to two stimuli) to be discriminated. The model modeled include phase differences between pairs of bandspass filtered between pairs of sinusoidally amplitude modulated tones, discrimination accuracy modulation, and discrimination of transient signals differing only in phase (response). © 2002 Acoustical Society of America.

43.65.Nn, 43.64.Bh (NfV)

most, both models do a good job of accounting for the effect of resolvability on listeners' sensitivity to phase differences between partials of a complex tone. A problem for models that explicitly discard across-channel timing information is that, for some stimuli, listeners can make across-channel phase comparisons. For example, when two groups of higher-numbered harmonics are filtered into separate frequency regions and presented concurrently, listeners can detect small (1-2 ms) shifts in the relative timing of the two groups, even when the outputs of auditory filters that might respond to both groups are masked by noise (Carlyon, 1994; Carlyon and Shackleton, 1994). Another example of sensitivity to across-channel timing differences is that subjects can detect a 30°-60° phase difference between the envelopes of two AM sinusoids, quite when the stimuli are presented against a noise background (Studdert et al., 1989; Yost and Shuff, 1989). Listeners' ability to use this type of information is also reflected in the phenomenon of co-modulation masking release (Soll et al., 1984), the size of which depends on the phase relationship between the envelopes of the separate noise bands that constitute the masker (Giguere et al., 1985; McFadden, 1986; Moore and Sjöström, 1990). Because "autocorrelation" and "auditory image" models analyze timing information on a within-channel basis, they fail to capture these across-channel differences. We will return to this point in more detail in Sec. IV.

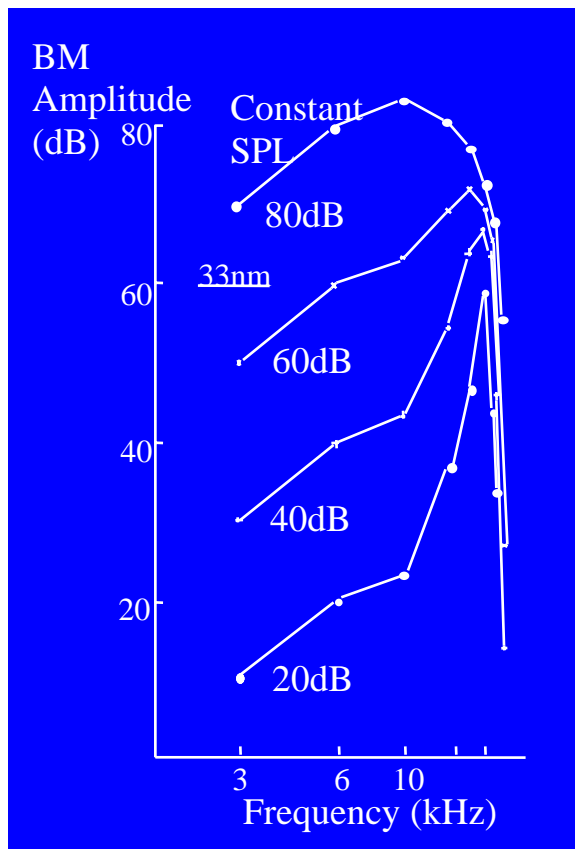
It seems clear, then, that a complete account of monaural phase sensitivity should not only explain why subjects cannot detect phase differences between resolved partials of complex tones, but also why they can detect across-frequency differences between the envelopes of broadband stimuli. In addition, such an account should be consistent

¹Carlyon and Shamma are now at the University of Cambridge.

R. Carlyon, S. Shamma. "An account of monaural phase perception." 2003.

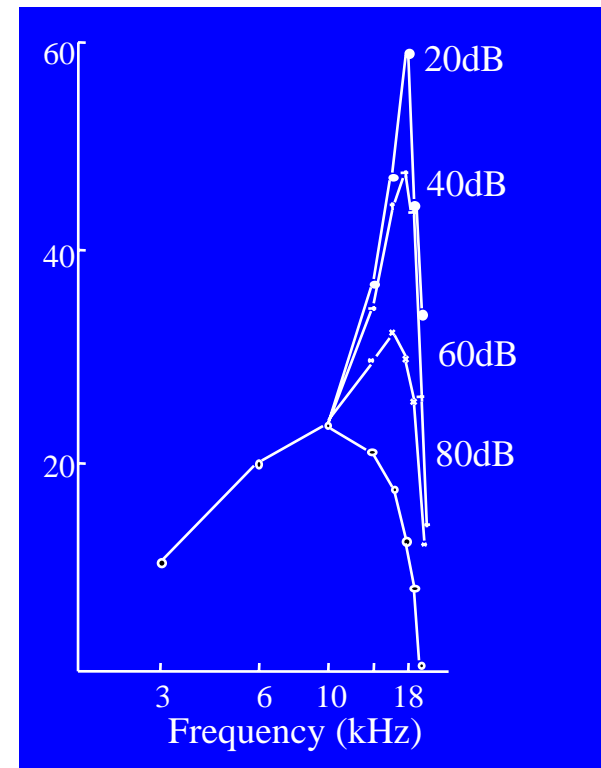
Bad Shihab

Nonlinear Cochlear Mechanics



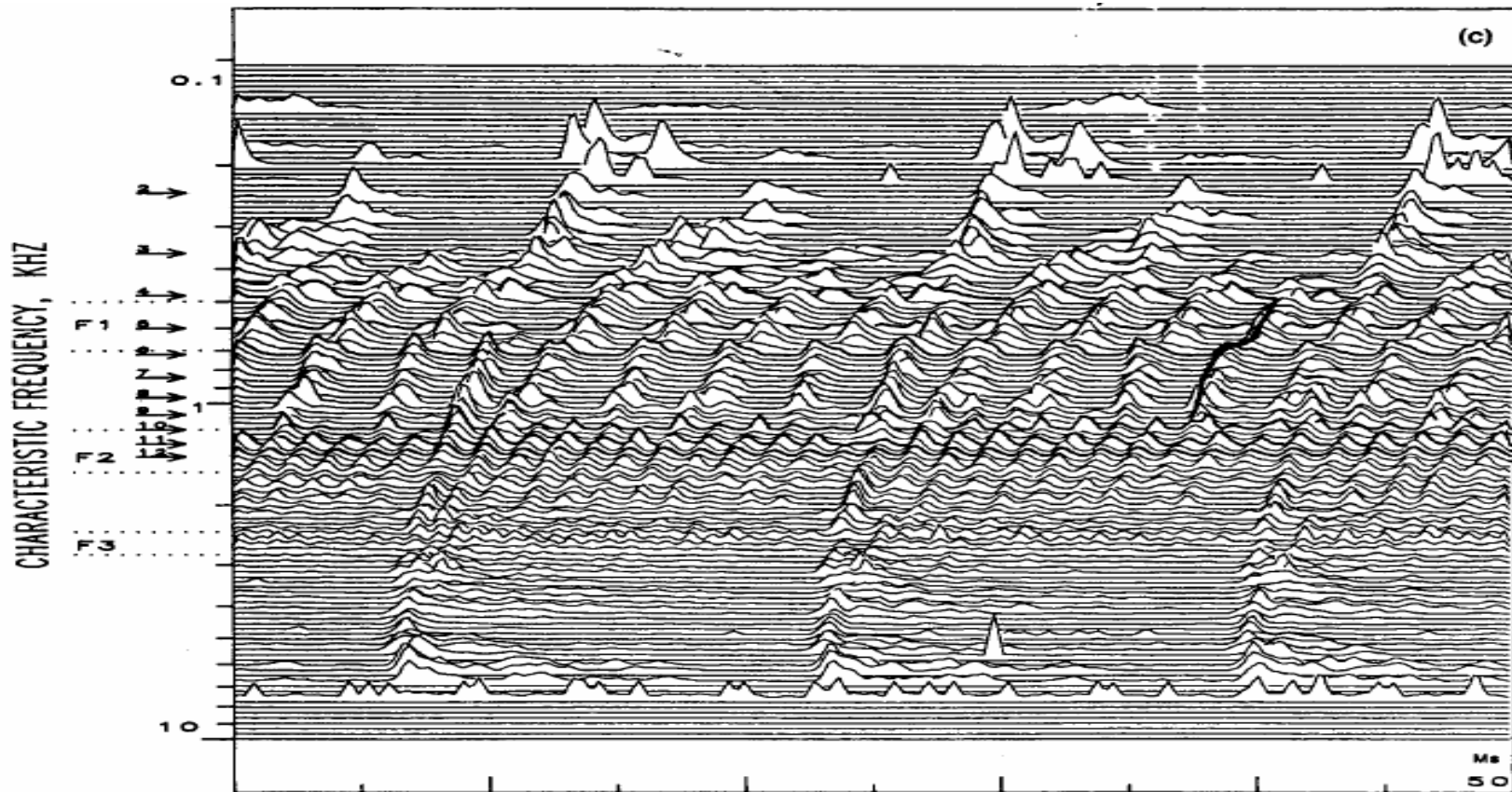
Mössbauer Data

SPL	Q_3	CF
80dB	1	10k
60dB	2.7	16k
40dB	4.8	17k
20dB	8.3	17k



From “Basilar membrane measurements and the travelling wave,” by B.M. Johnstone, R. Patuzzi, and G. K. Yates, *Hearing Research*, 22 (1986) 147-153.

Cat Vowels (/da/)





Goal

- Test phase perception models with more sophisticated cochlear models
- Auditory models
 - Gammatone
 - Gammatone with hair cell model
 - Lyon's passive Ear
 - STRF



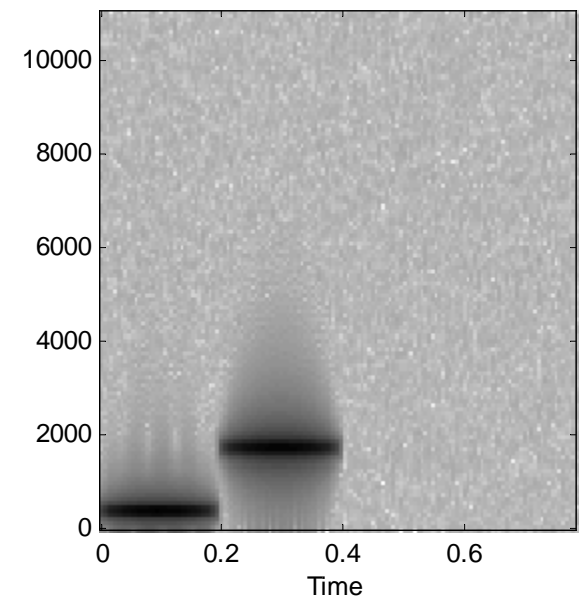
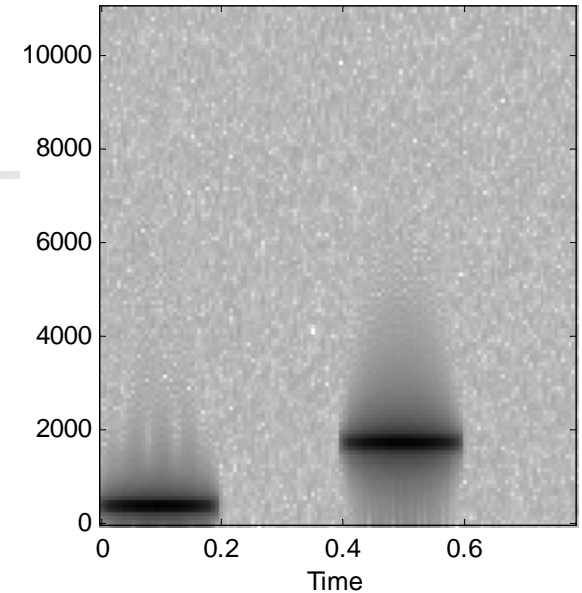
Nonlinearities

- Capture effect
 - Dominant frequency is represented
 - A few dB matter
- Stochastic resonance
 - Sub-threshold signals entrain firings
- Temporal coding
 - Non-linear mapping

Phase Models

- Within “channel”
 - Auditory nerve channel
 - Phase matters

- Between channels
 - Phase does matter!
 - Single sinusoids
 - 2 octave separation





Phase Perception Model

- Short-term

- Within channel
- Sensitive to small phase change
- STRF fails



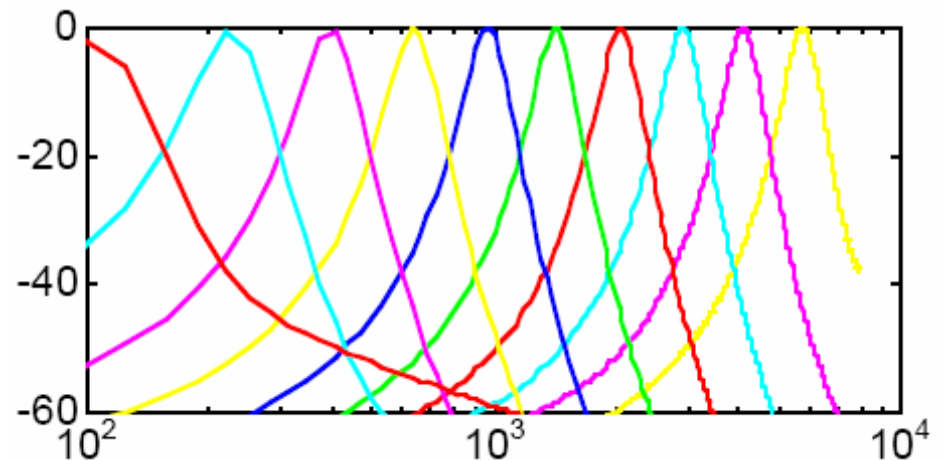
Today

- Long term

- Speech-rate perception
- Large phase changes (relative to CF)
- STRF works great

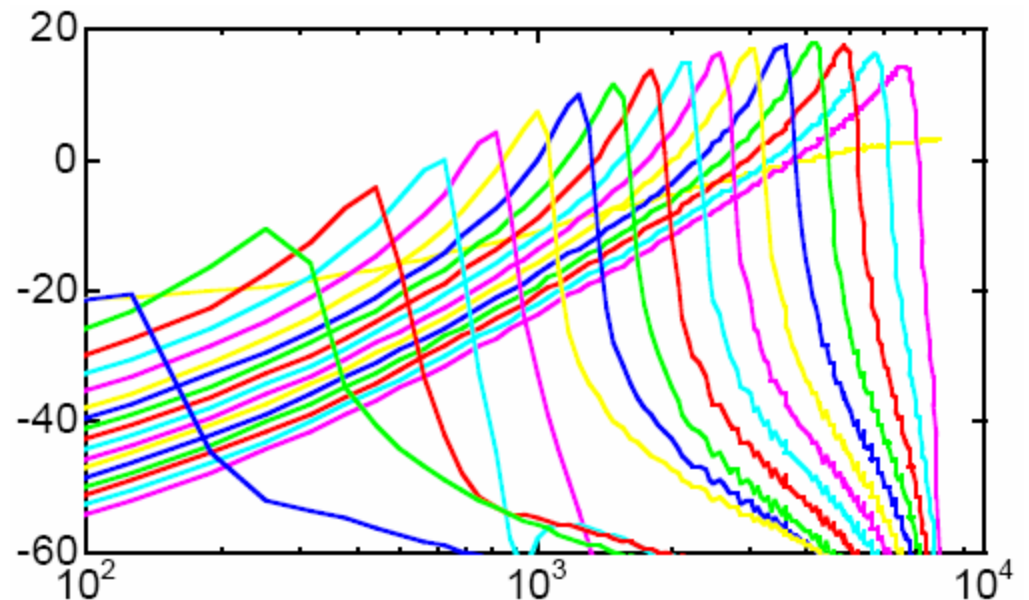
Models – Gammatone

- Critical band filters
 - Psychoacoustic data
 - Linear bandpass filters
- Meddis Hair Cell
 - Nonlinear reservoir model



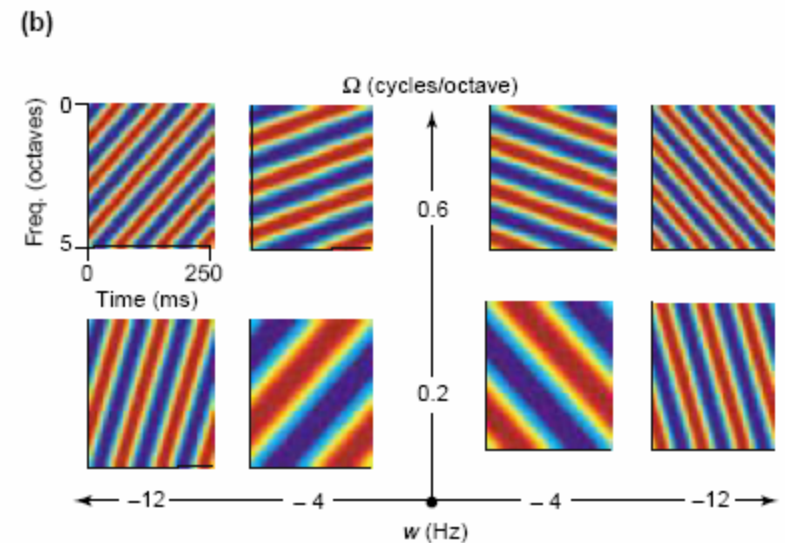
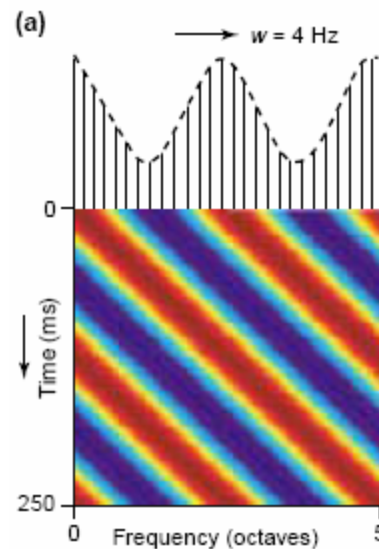
Models – Lyon's Passive Ear

- Lowpass filters
 - Transmission line
 - Cascade
- Nonlinear automatic gain control (AGC)
 - Passive



Models – STRF

- Spatial temporal response function
 - Based on Shamma's cortical data
- Fit spectrogram data
 - Two-dimensional sinusoids
 - Image domain



Patterson Phase Results

- Stimulus
 - Constant phase shift per critical band
 - Shifted phase gradually
 - Small local changes
 - Large global change
- Result
 - Couldn't prove channel effect!

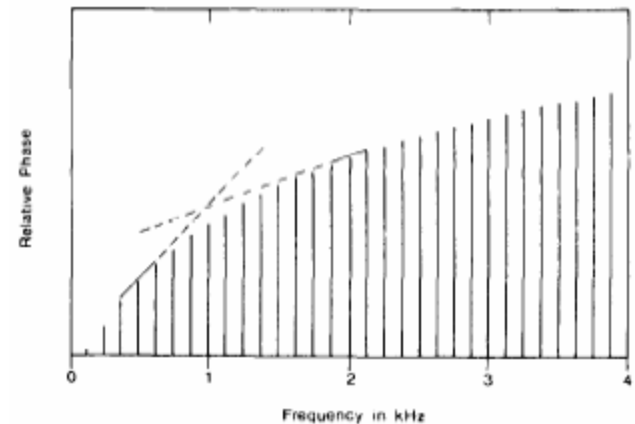
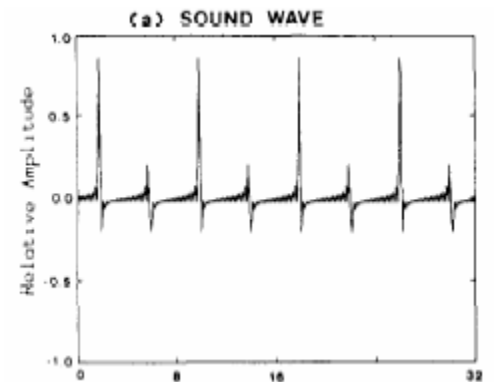
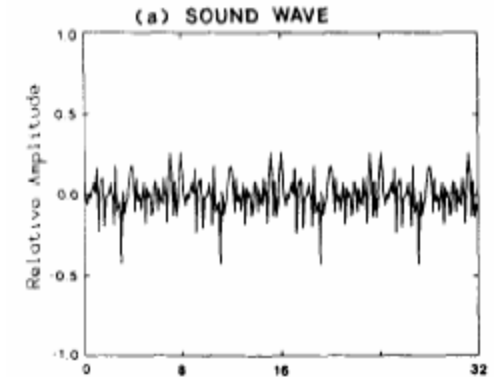
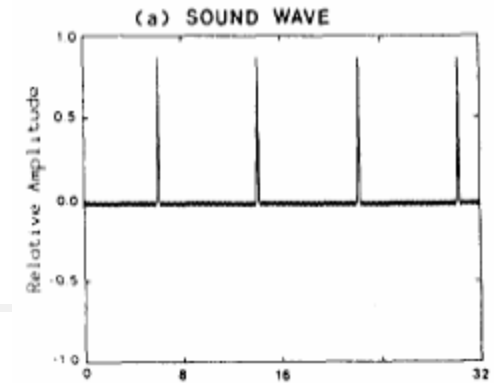


FIG. 11. The general form of the phase spectrum for the MPH waves. Since the function is approximately linear in any local region, it minimizes WC phase changes. At the same time, the effect of the deceleration cumulates across larger frequency regions to produce significant BC phase shifts.

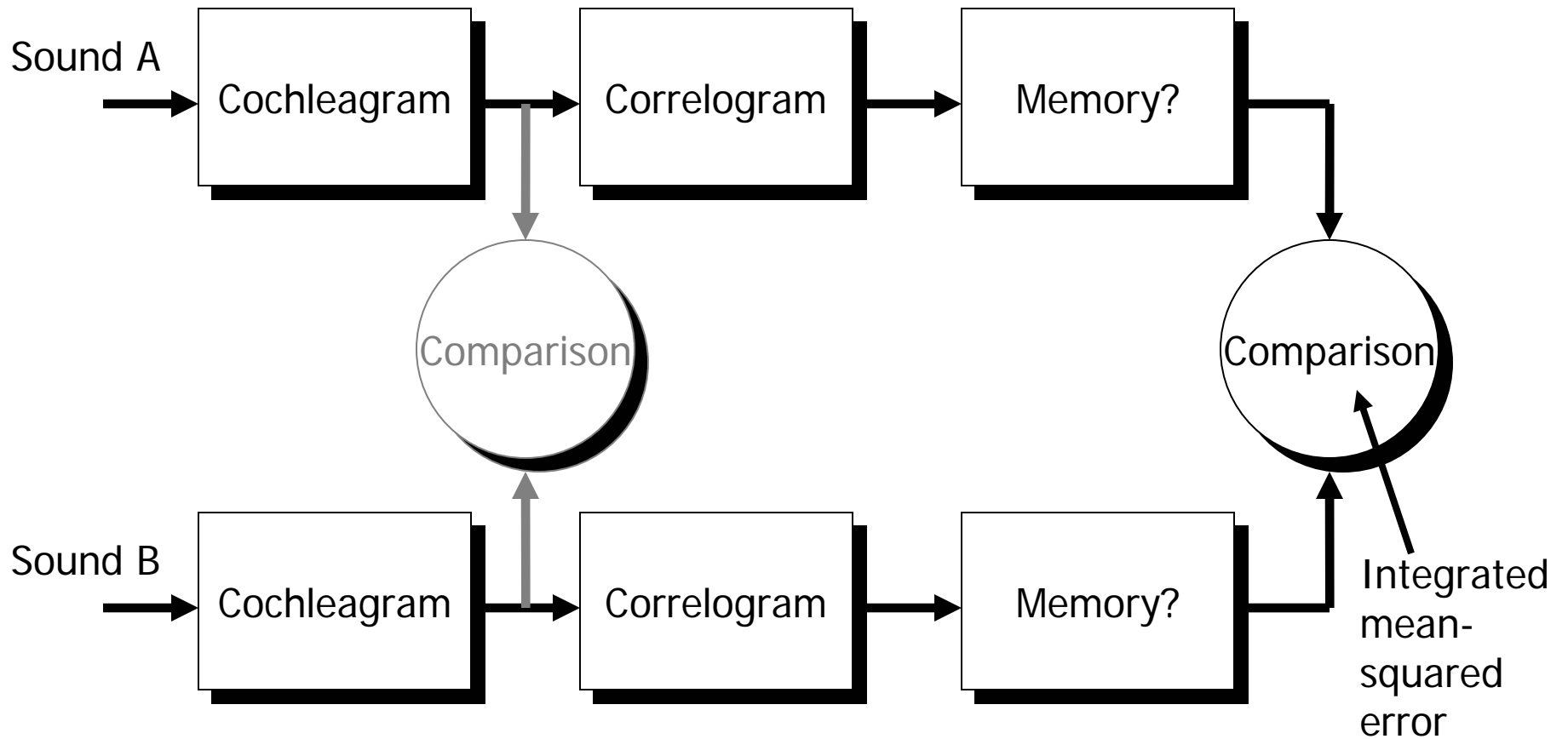
Patterson Stimuli

- Harmonic signal
 - 31 harmonics
- Phase Alignment
 - Cosine phase
 - Random phase
 - Alternating phase (APH)



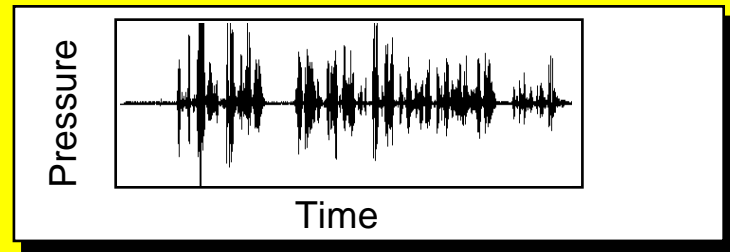


Our Model



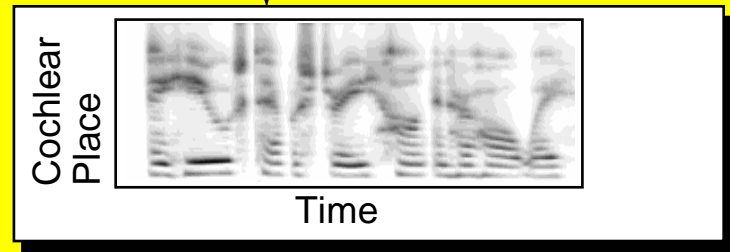
Auditory Dimensions

One Dimensional



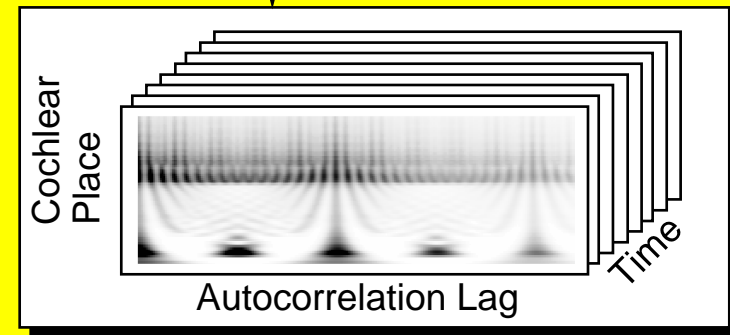
Cochlear Processing

Two Dimensional

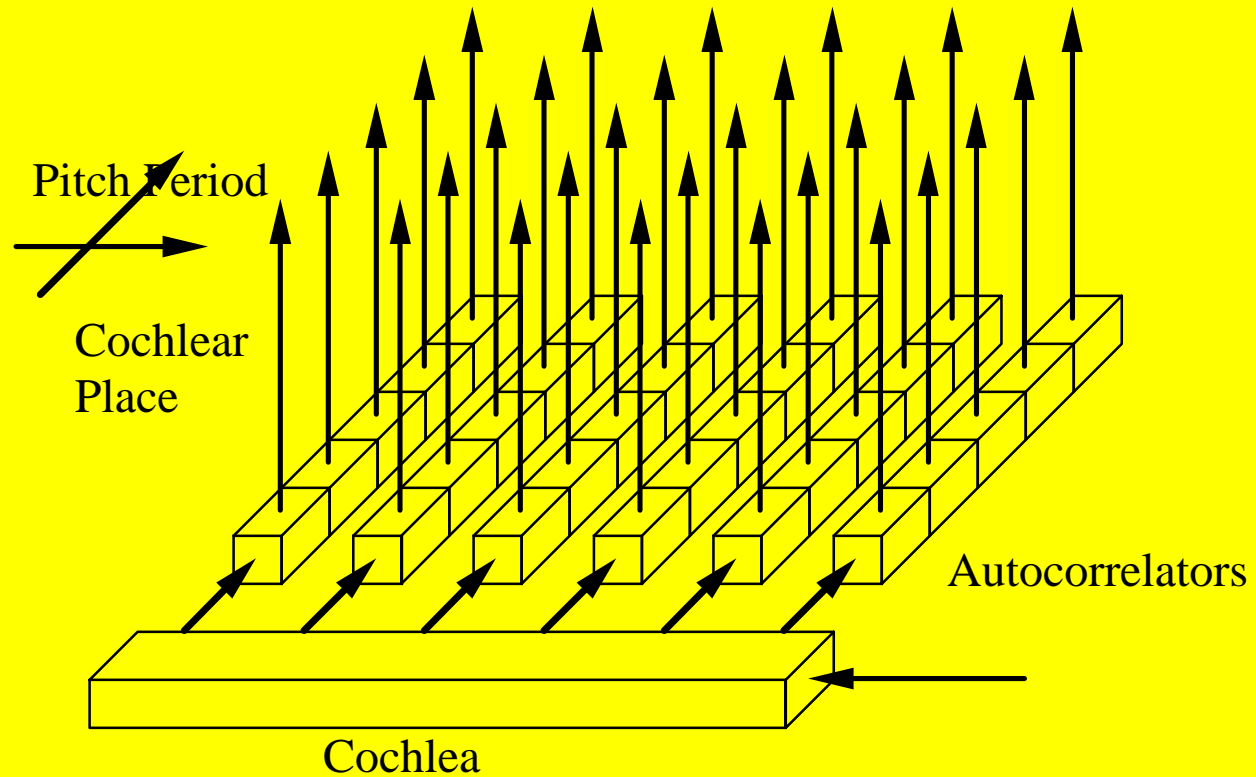


Correlogram Processing

Three Dimensional



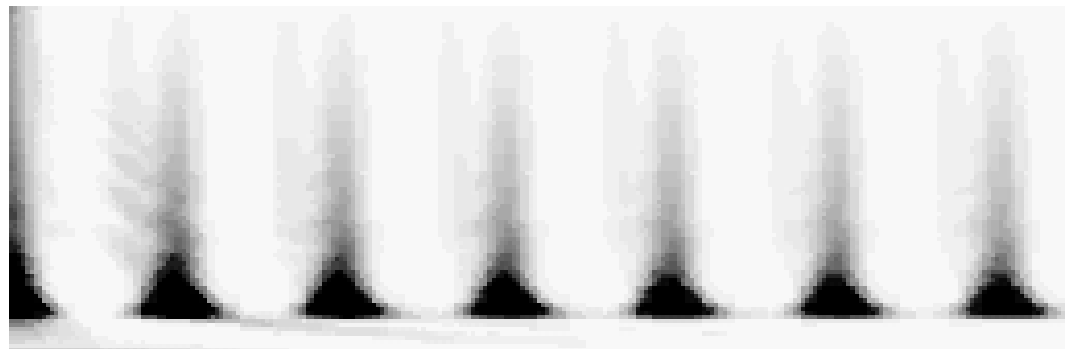
Correlogram Structure





Duda Tones

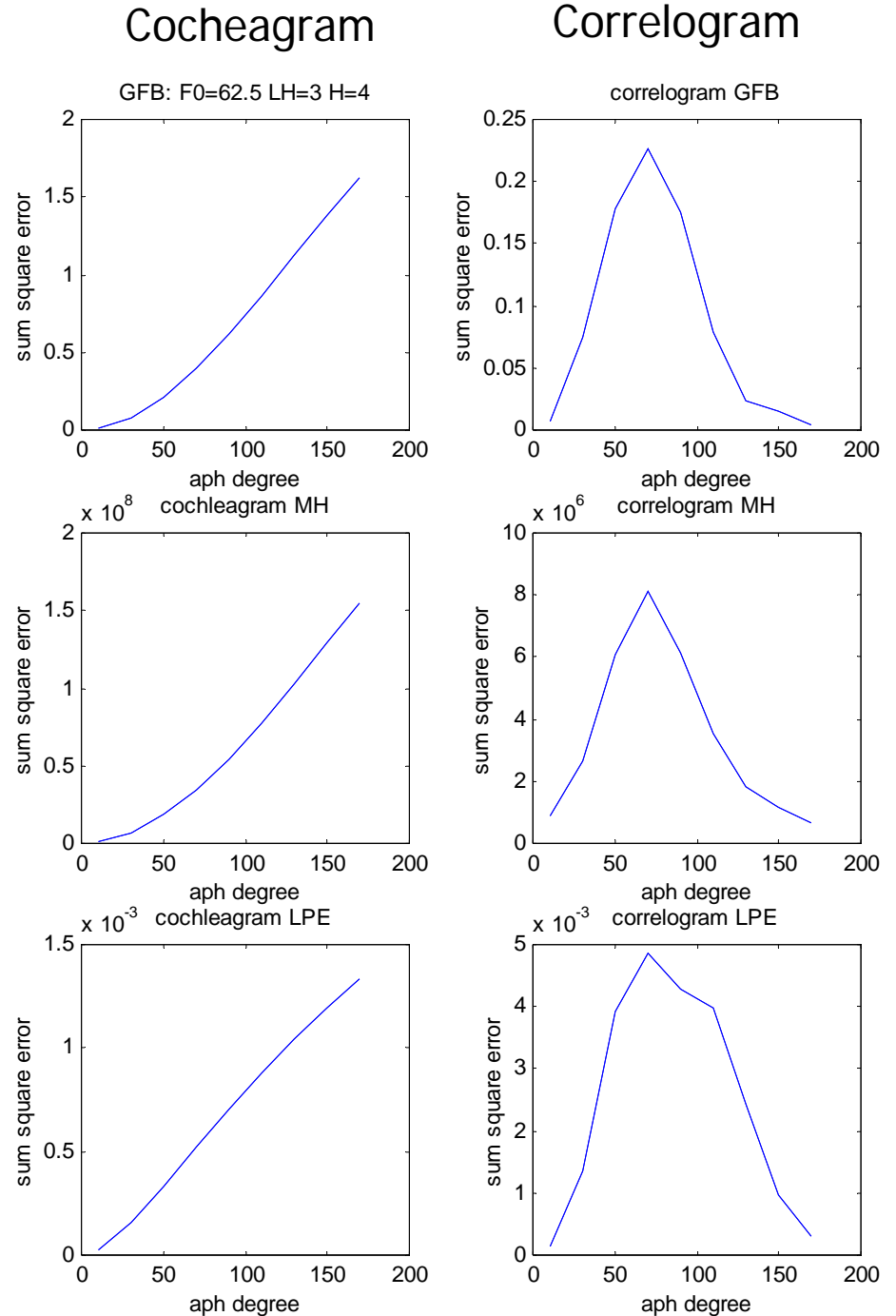
- Simple tone
- Harmonic complex
- Tone build up (slow and fast)
- Tone removal (from top and bottom)
- Wiggle





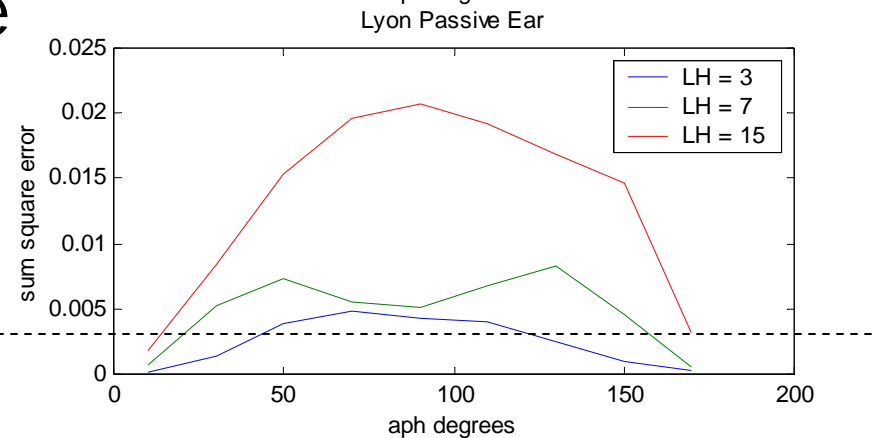
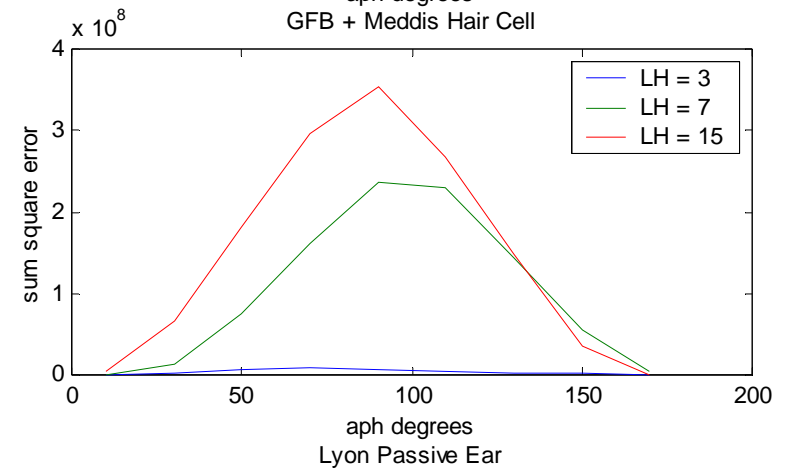
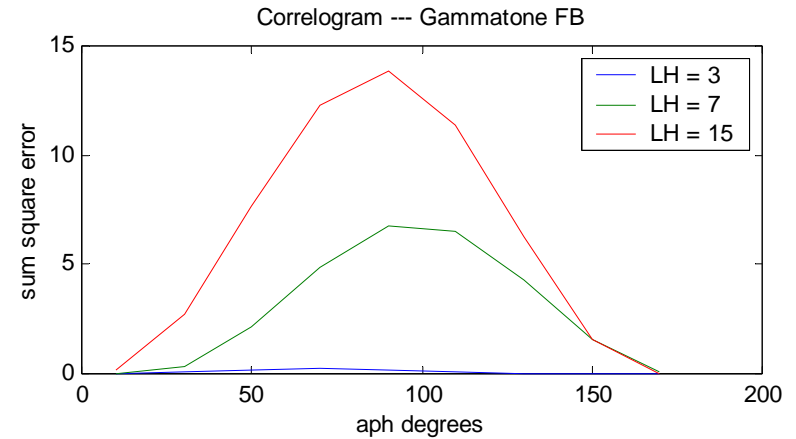
Results

- APH
 - Alternating phases are shifted
- Cochleagram
 - Auditory nerve
 - Error grows
- Temporal pattern
 - Error falls



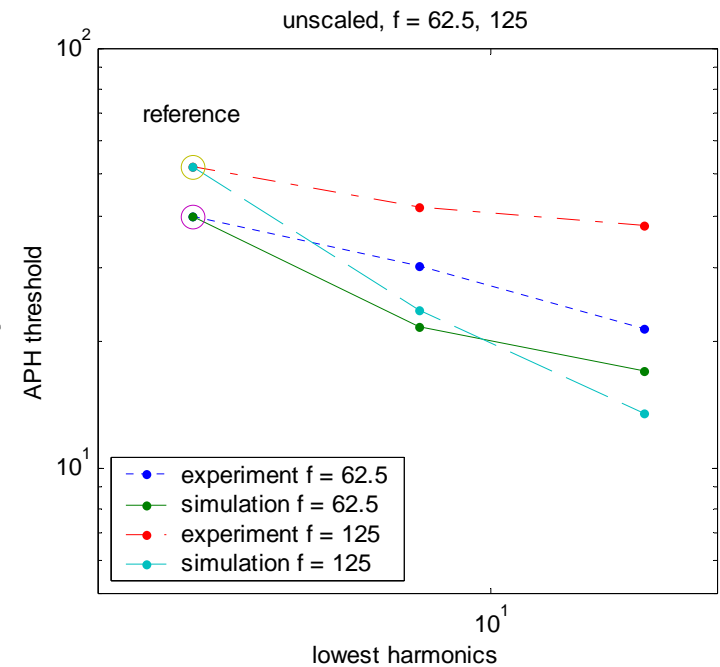
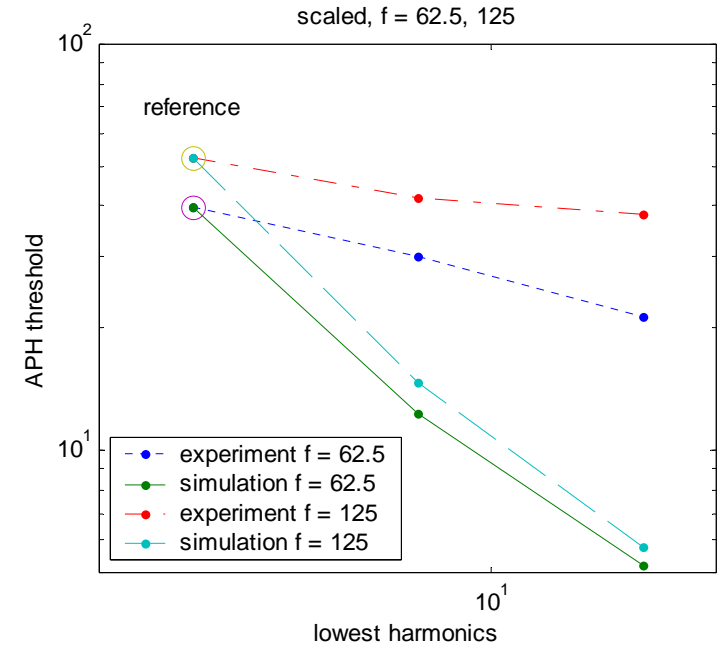
APH Results

- Test
 - Error wrt phase
 - Lowest harmonic
- Gammatone
 - Too much variance
 - No one threshold
- Lyon's Ear



Threshold

- Gammatone
 - No one threshold
- Lyon's Passive Ear
 - Consistent results
 - Depends on loudness (scaling)



Craig and Jeffress

- Two components
- Reverse polarity
- Task
 - Detect difference
 - Identify phase

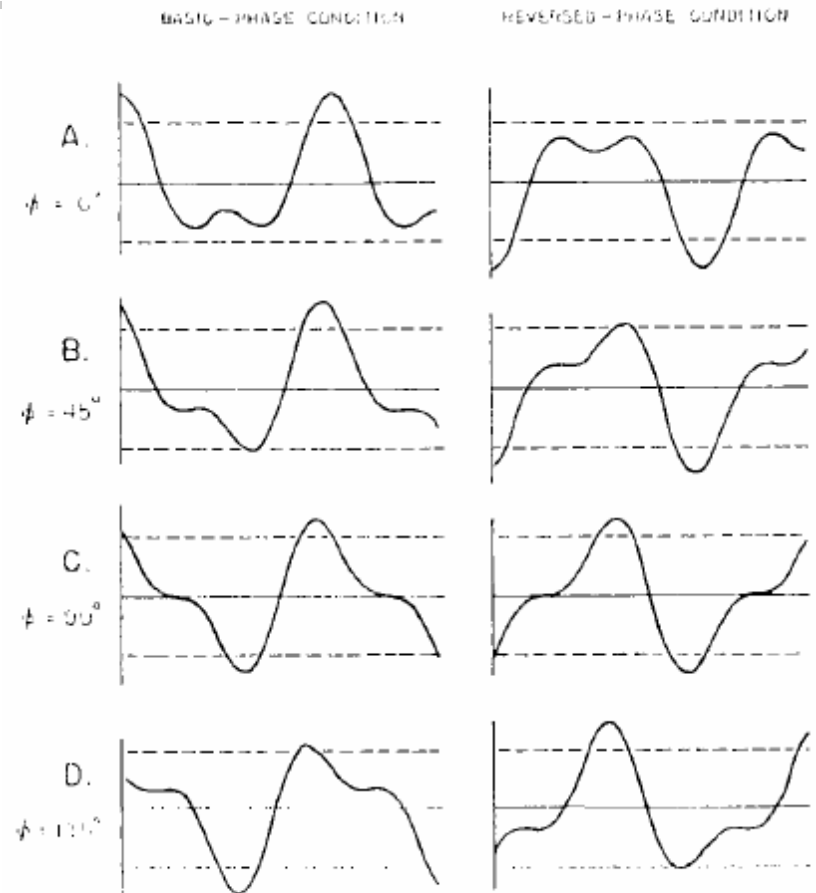
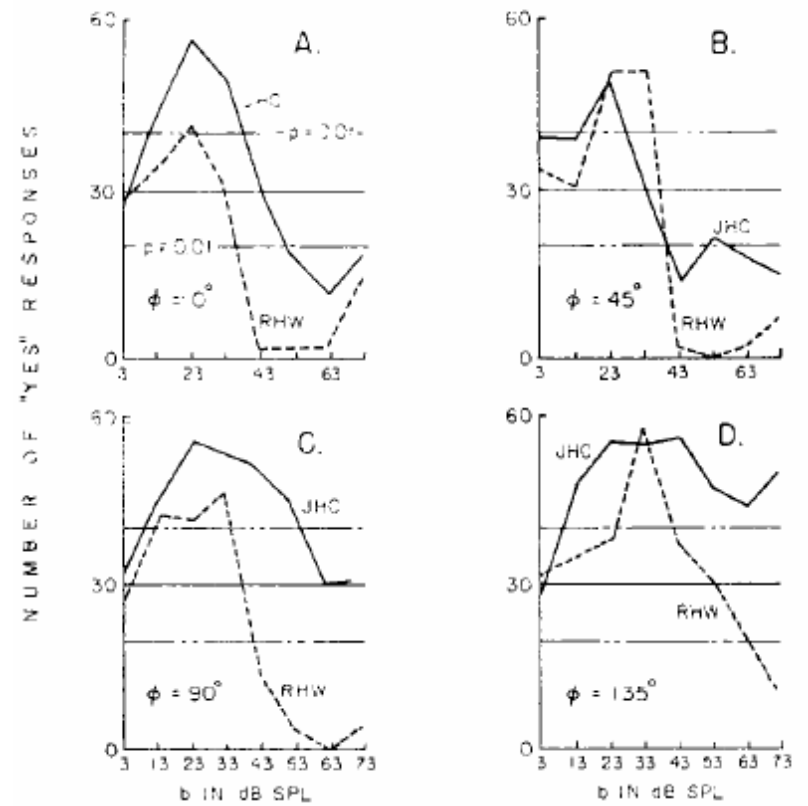
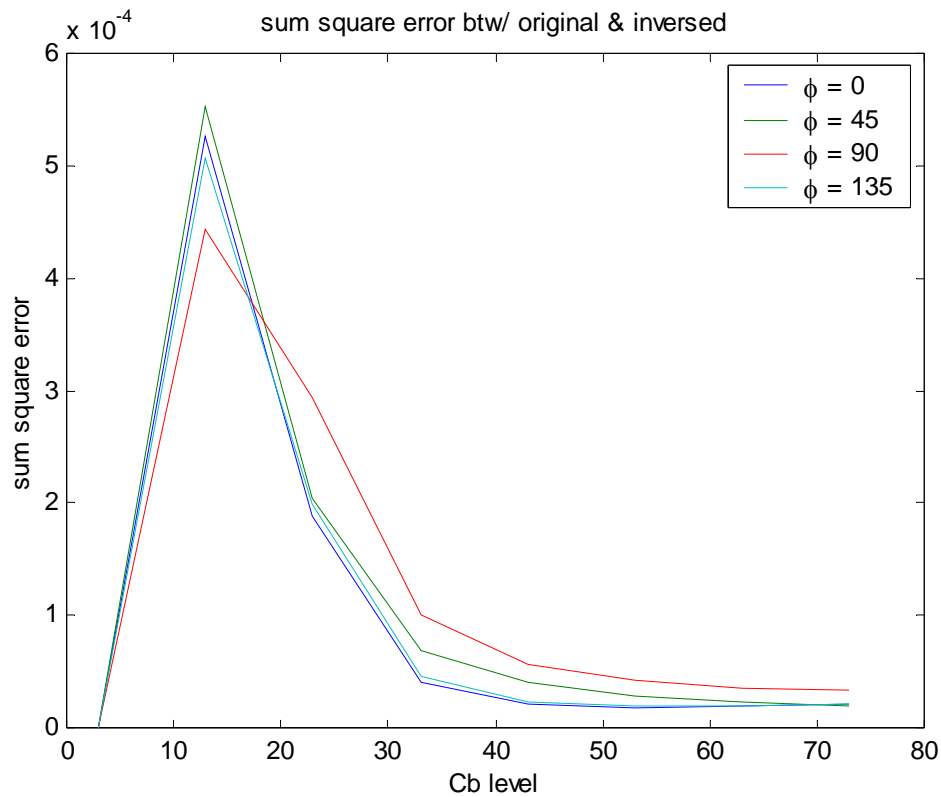


FIG. 3. The two-component stimulus, $Y = A \cos \omega t + B \cos (2\omega t + \phi)$, for the four phase conditions used in this study (drawn with an arbitrarily selected component-amplitude ratio, $A = 2B$). Positive values indicate inward displacements of the eardrum.

Craig & Jeffress Results





To Do

- Improve results
 - Tune model bandwidths
 - Better cochlear/hair cell integration
- Verify two distinct regimes
 - Break in performance
 - Not

Conclusions

- Phase perception model
 - Two different models
 - Short-term within channel
 - Long-term STRF?
- Better auditory model
 - Less about critical bands

