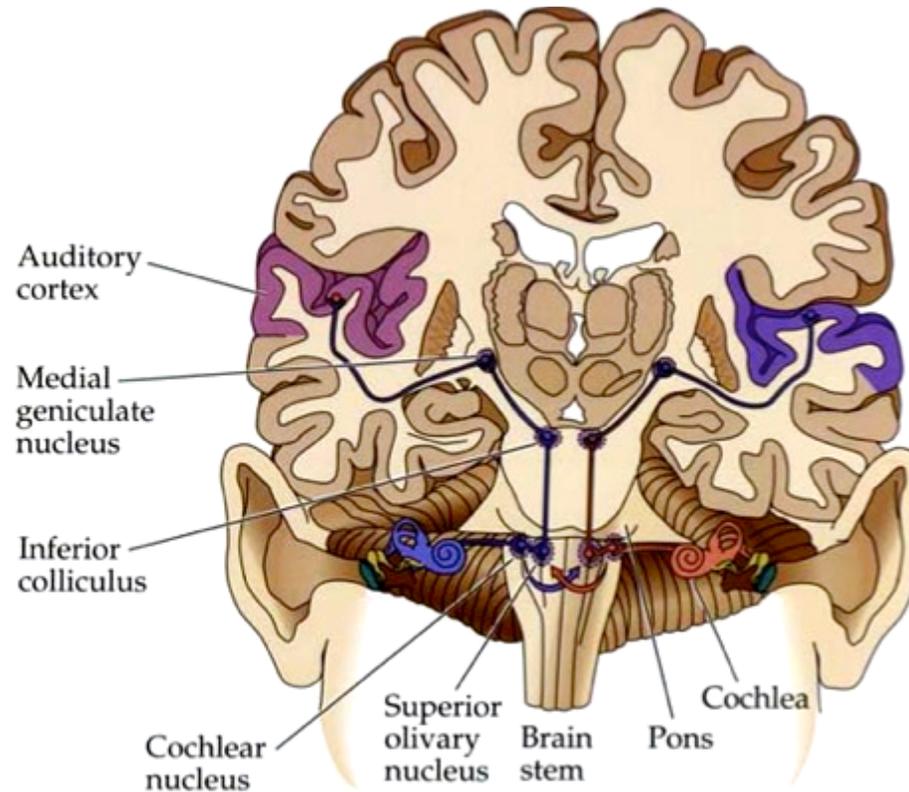
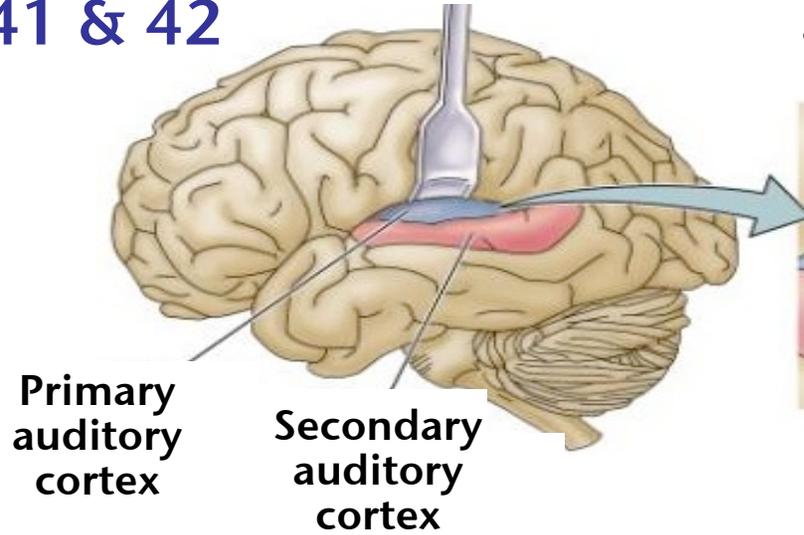


# Human Auditory Cortex

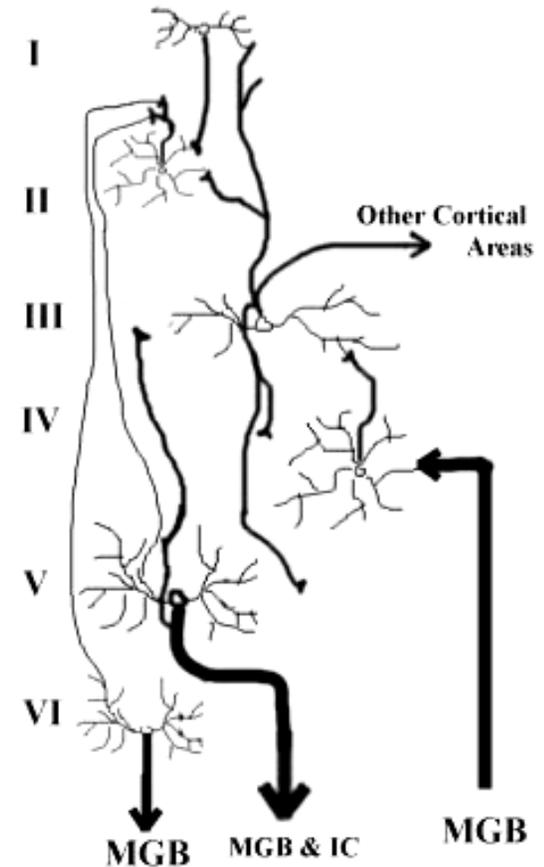
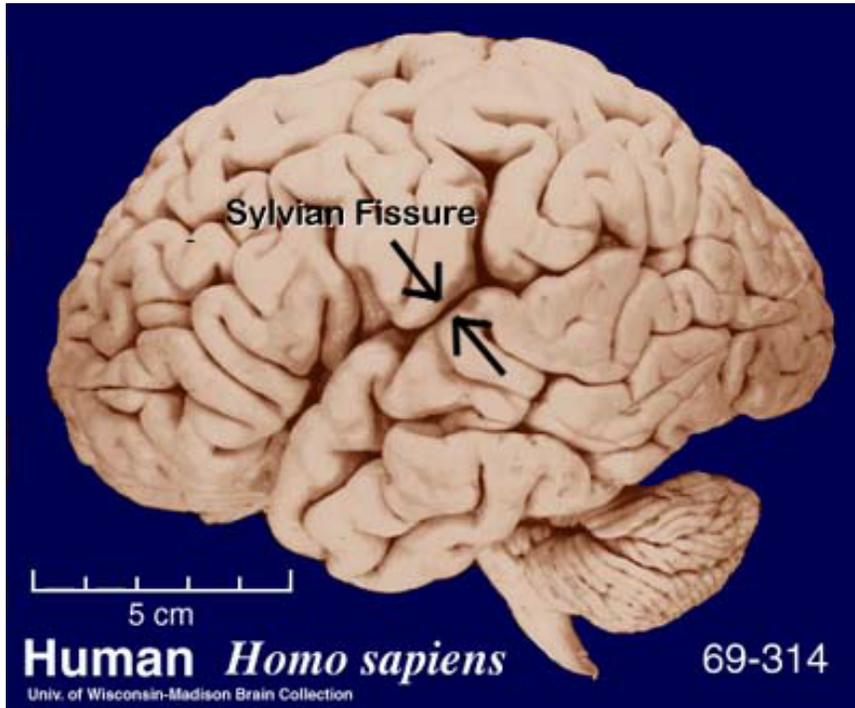
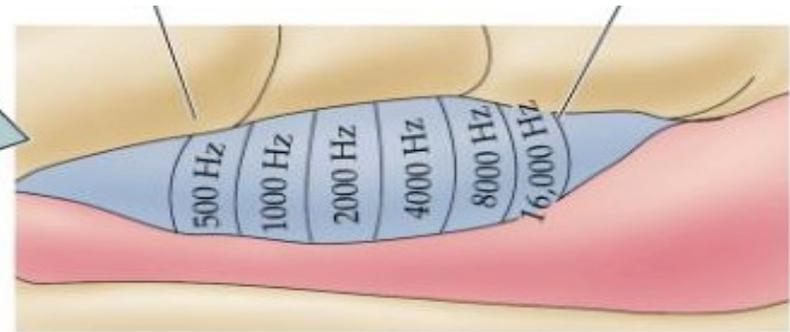


# Areas 41 & 42

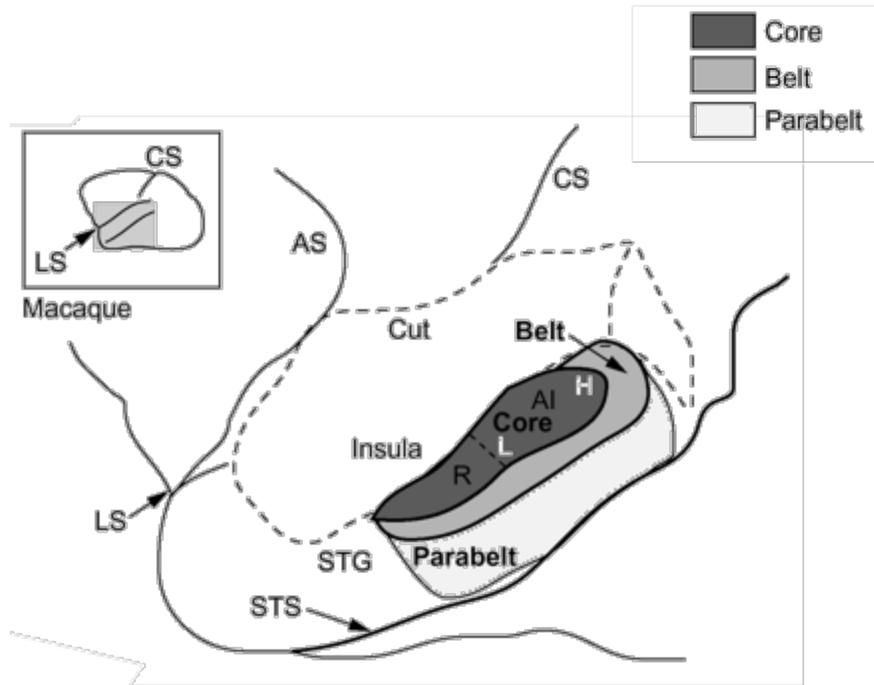


Corresponds to apex of cochlea

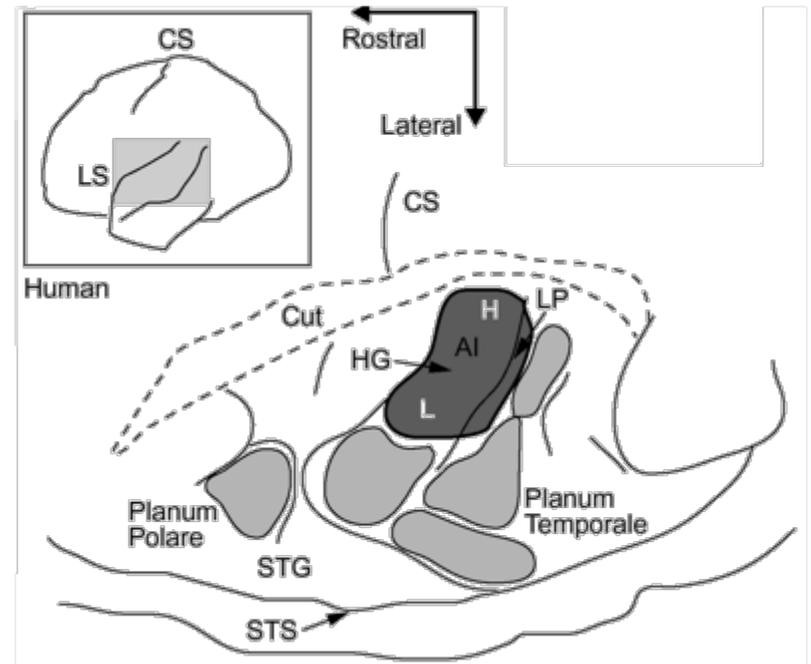
Corresponds to base of cochlea



# Auditory Cortex in Macaque & Human



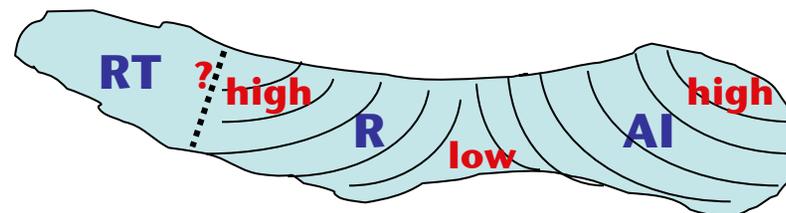
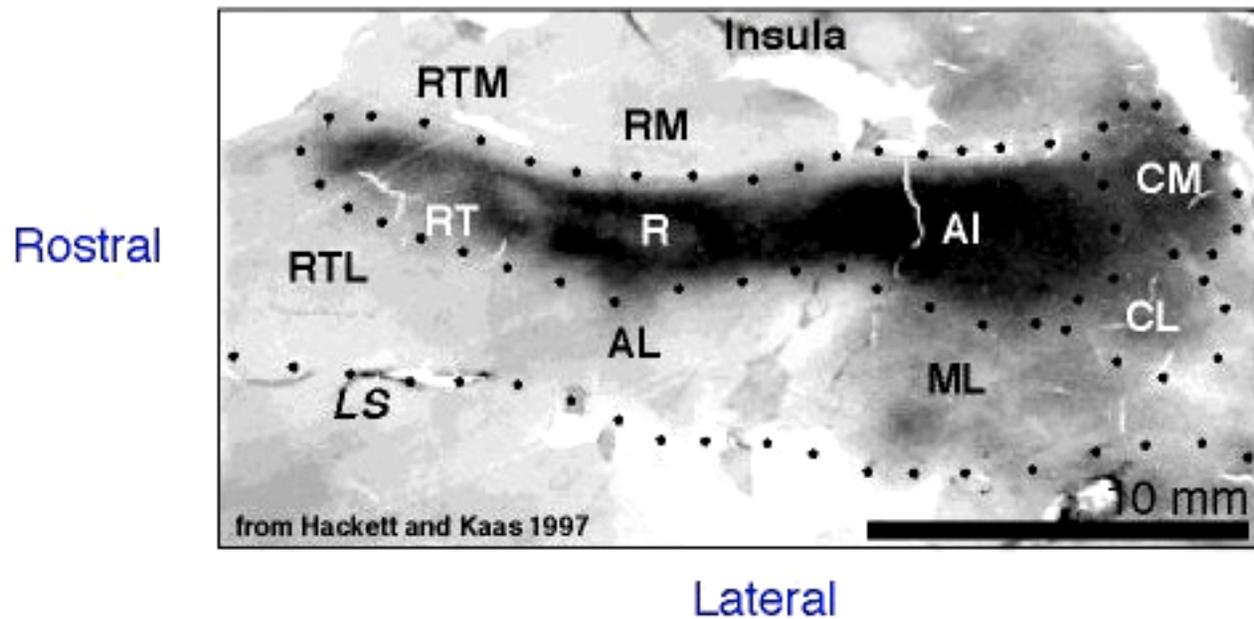
Macaque



Human

# Parvalbumin Defines Macaque Core Auditory Cortex

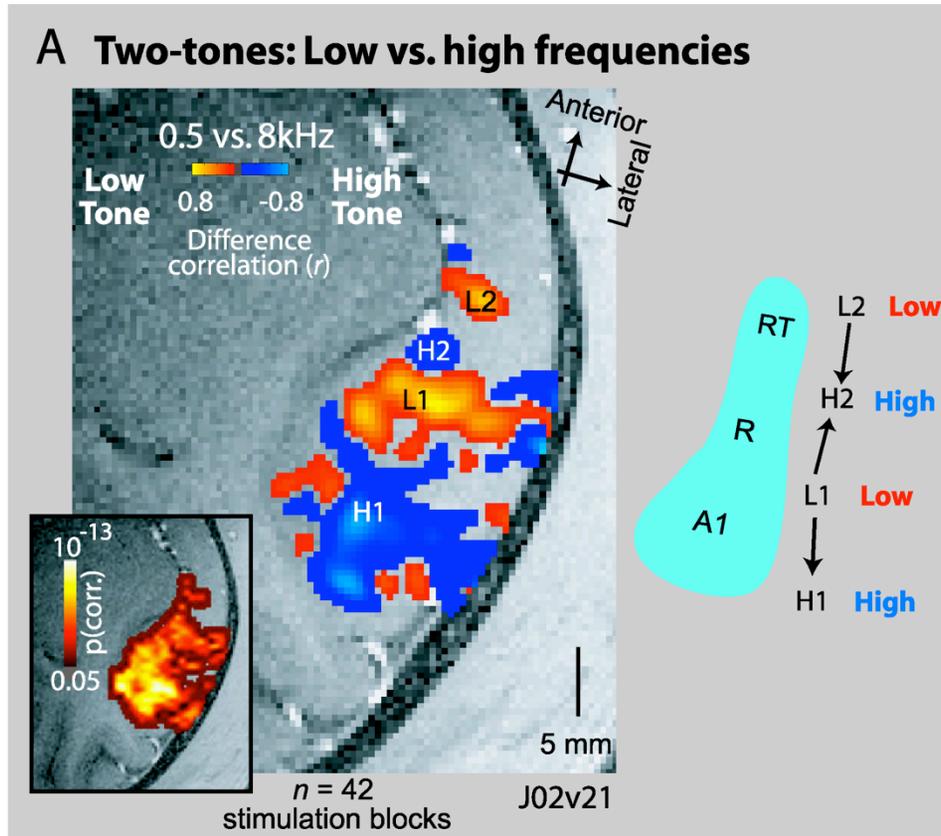
Flattened section of superior temporal plane



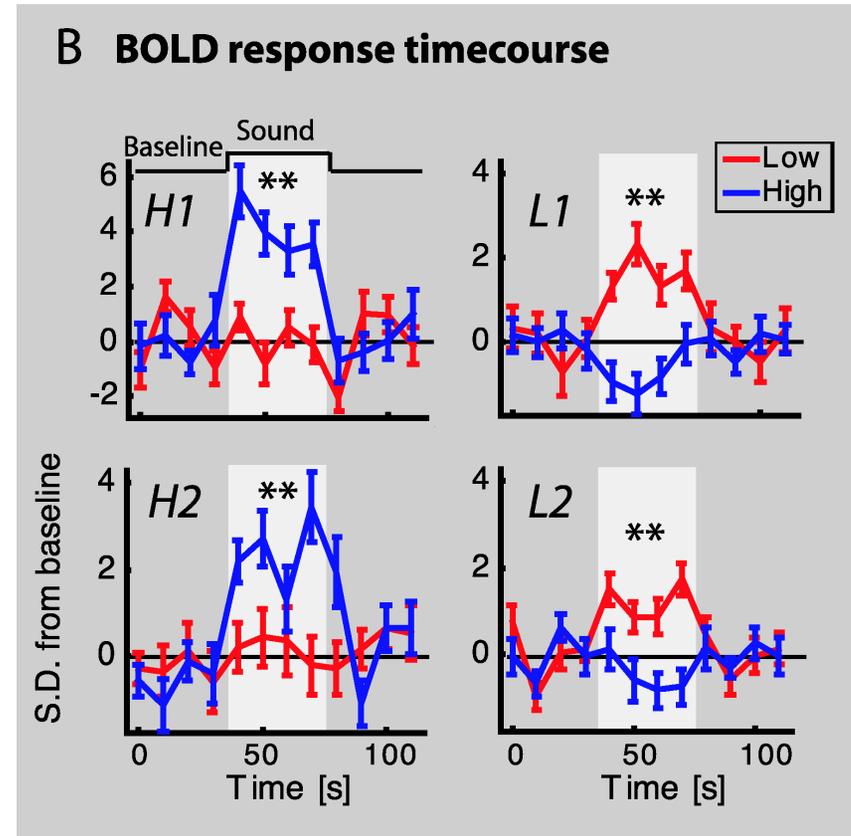


# Functional Imaging in Macaque Auditory Cortex

Petkov et al. 2006

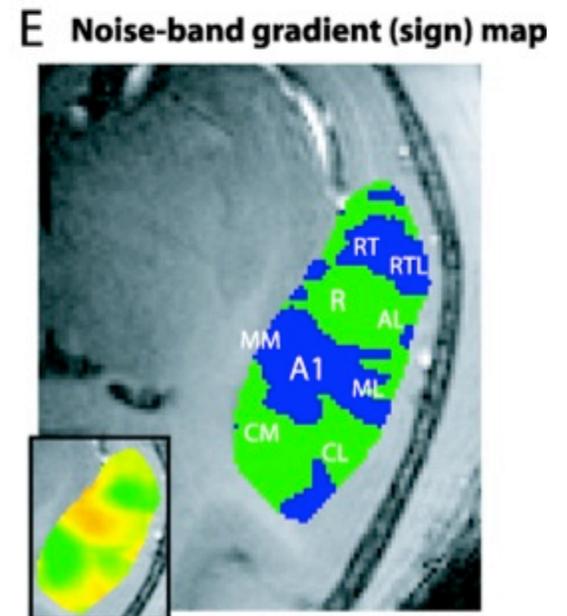
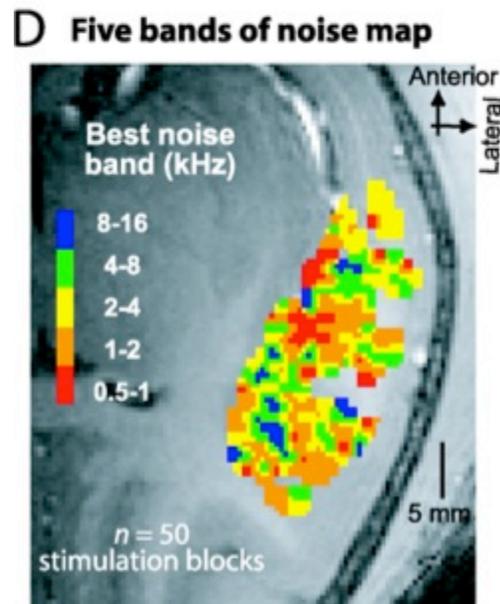
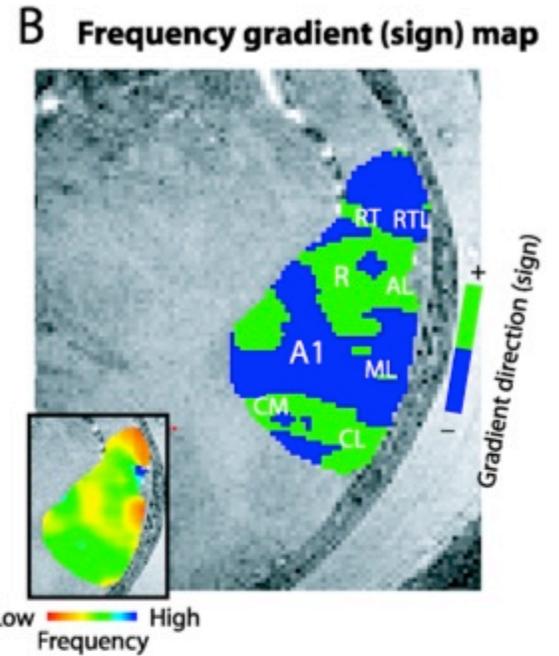
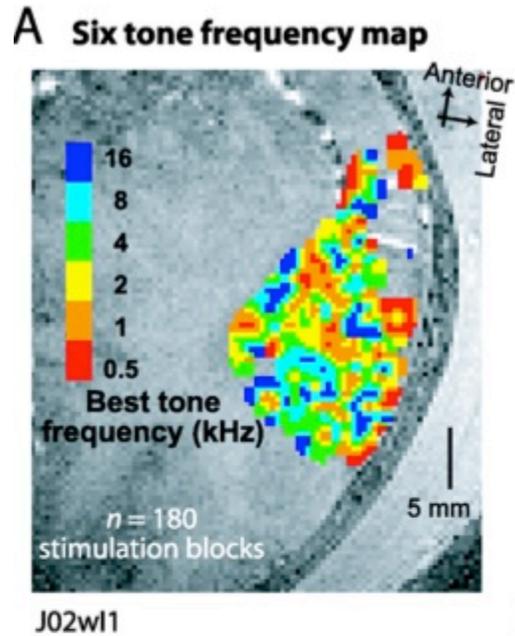
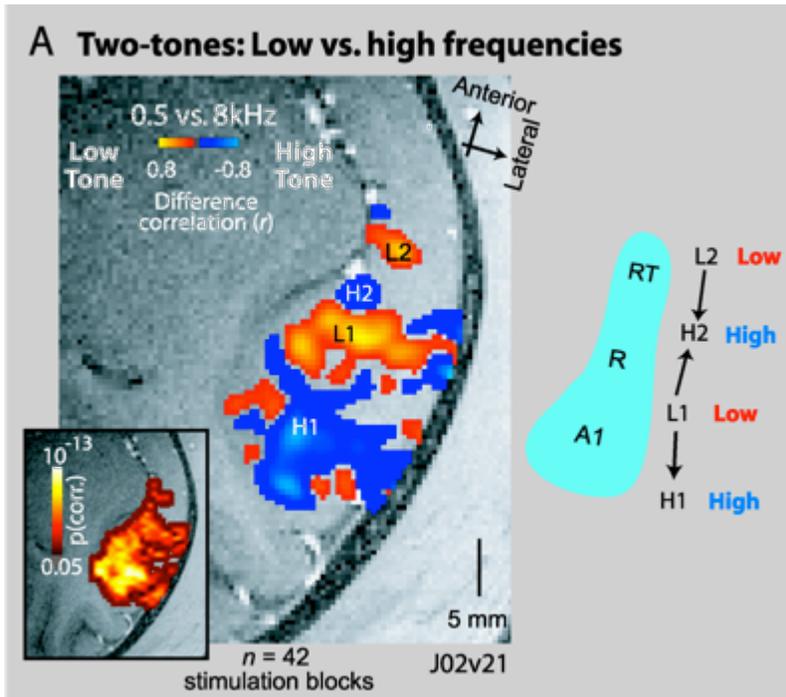


**Alternating blocks of 0.5 kHz and 8 kHz tones.**  
For significantly active voxels (inset), frequency selectivity = difference in signal correlation to Low vs. High tones. Results are thresholded to show only strongly frequency selective voxels.



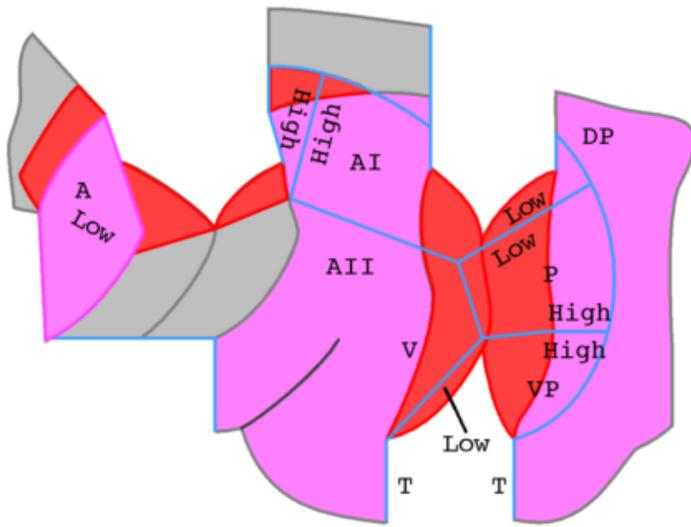
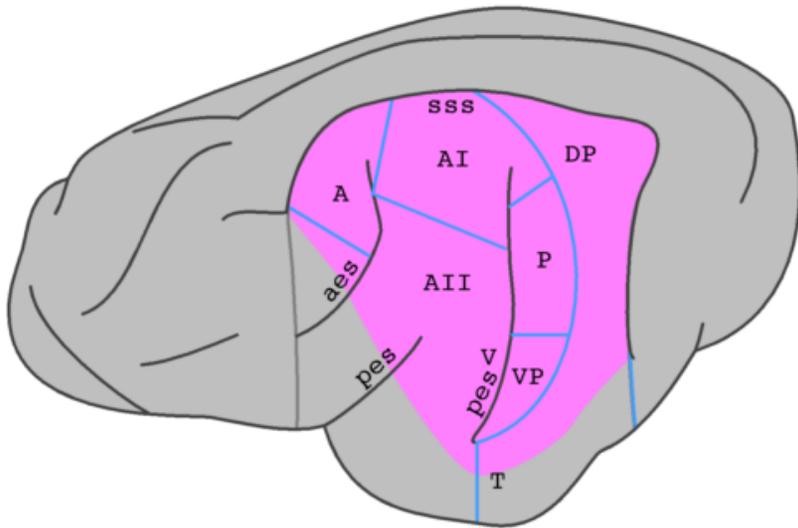
**Timecourse of BOLD response** illustrating that clusters of voxels underneath labels in (A) showed greater activation to the high frequency tone at H1 and H2 but to the low frequency tone at L1 and L2

# Functional Imaging in Macaque Auditory Cortex

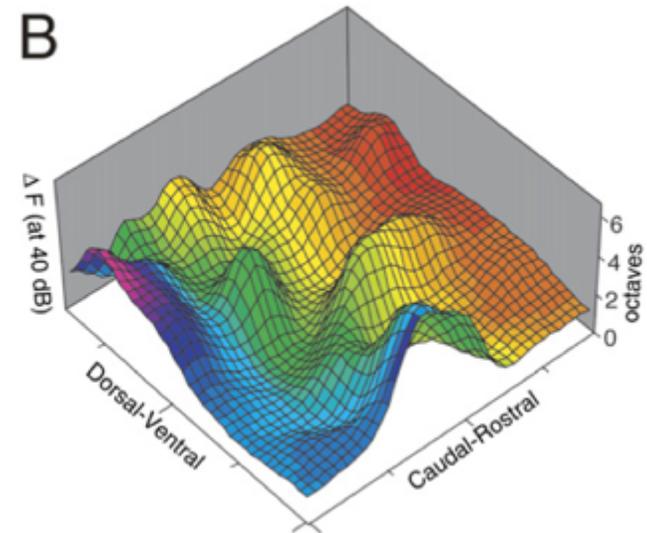
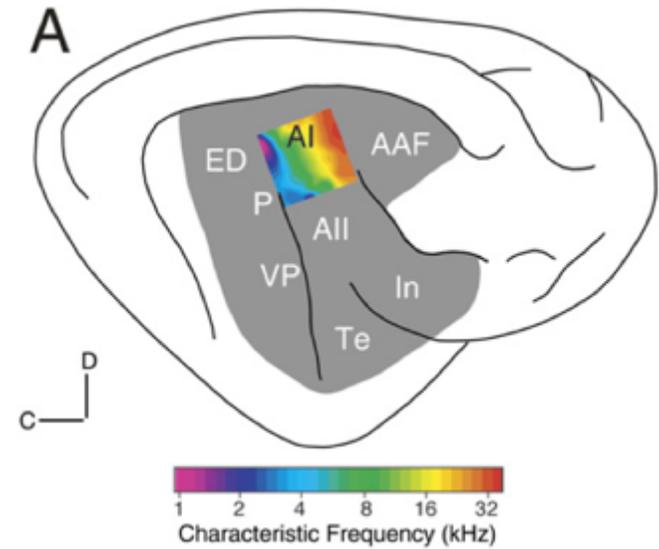


*Petkov et al. 2006*

# Multiple tonotopic fields in cat auditory cortex



Reale & Brugge (1990)



Schreiner & Weiner (2007)

# AI tonotopy studied via *in vivo* 2-photon microscopy

anesthetized mice, no correction for laser noise or acoustic calibration

## Functional organization and population dynamics in the mouse primary auditory cortex

Gideon Rothschild<sup>1,2</sup>, Israel Nelken<sup>1,2</sup> & Adi Mizrahi<sup>1,2</sup>

*Nat Neurosci* 2010, 13:353-360

sound source 10cm from  
contra ear

50ms tones,  
5 ms rise/ fall  
8 repetitions per stim

Cortical processing of auditory stimuli involves large populations of neurons with distinct individual response profiles. However, the functional organization and dynamics of local populations in the auditory cortex have remained largely unknown. Using *in vivo* two-photon calcium imaging, we examined the response profiles and network dynamics of layer 2/3 neurons in the primary auditory cortex (A1) of mice in response to pure tones. We found that local populations in A1 were highly heterogeneous in the large-scale tonotopic organization. Despite the spatial heterogeneity, the tendency of neurons to respond together (measured as noise correlation) was high on average. This functional organization and high levels of noise correlations are consistent with the existence of partially overlapping cortical subnetworks. Our findings may account for apparent discrepancies between ordered large-scale organization and local heterogeneity.

## Dichotomy of functional organization in the mouse auditory cortex

Sharba Bandyopadhyay<sup>1,2</sup>, Shihab A Shamma<sup>2,3</sup> & Patrick O Kanold<sup>1,2</sup>

*Nat Neurosci* 2010, 13:361-368.

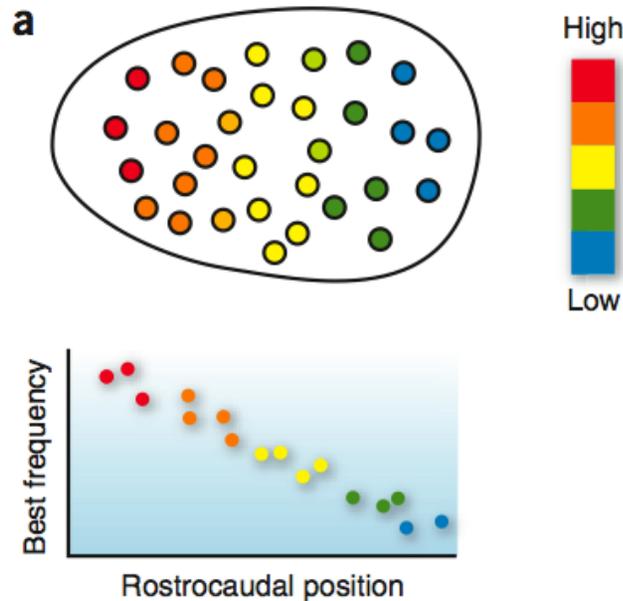
sound source possibly  
coupled to contra ear (not  
specified clearly)

The sensory areas of the cerebral cortex possess multiple topographic representations of sensory dimensions. The gradient of frequency selectivity (tonotopy) is the dominant organizational feature in the primary auditory cortex, whereas other feature-based organizations are less well established. We probed the topographic organization of the mouse auditory cortex at the single-cell level using *in vivo* two-photon Ca<sup>2+</sup> imaging. Tonotopy was present on a large scale but was fractured on a fine scale. Intensity tuning, which is important in level-invariant representation, was observed in individual cells, but was not topographically organized. The presence or near absence of putative subthreshold responses revealed a dichotomy in topographic organization. Inclusion of subthreshold responses revealed a topographic clustering of neurons with similar response properties, whereas such clustering was absent in supra-threshold responses. This dichotomy indicates that groups of nearby neurons with locally shared inputs can perform independent parallel computations in the auditory cortex.

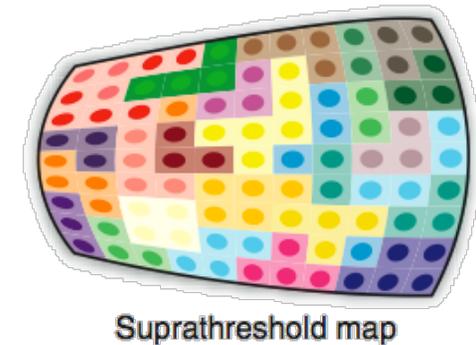
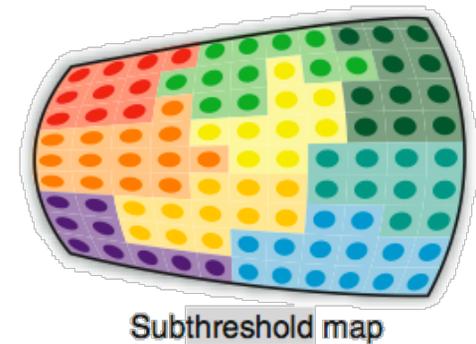
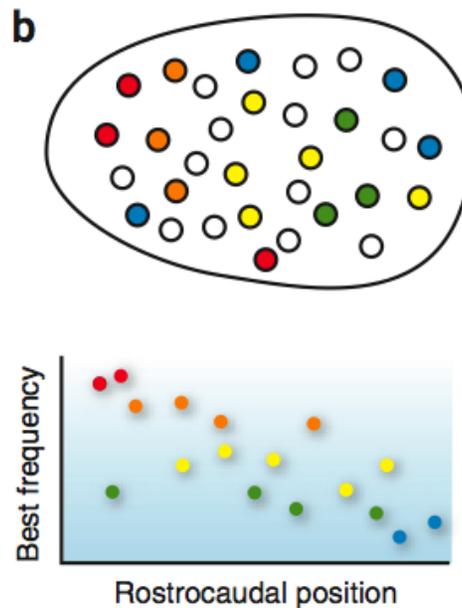
# Dichotomy in mouse AI tonotopy

*in vivo* two-photon  $\text{Ca}^{2+}$  imaging

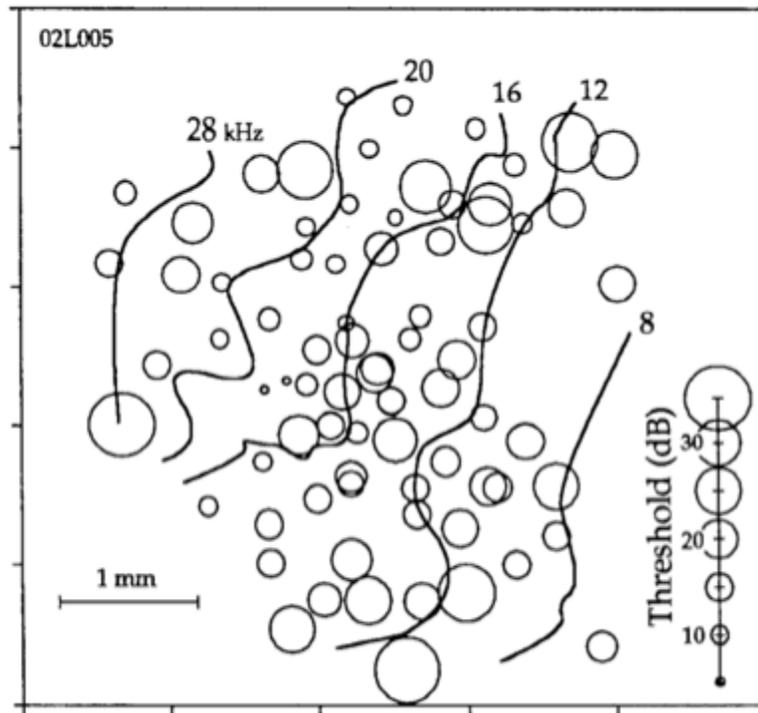
Classical tonotopic map shows a smooth increase in the preferred frequency along a rostrocaudal axis.



Two-photon imaging reveals tonotopy only on a large scale.  
*(open circles: no clear BF, or unresponsive to tones.)*



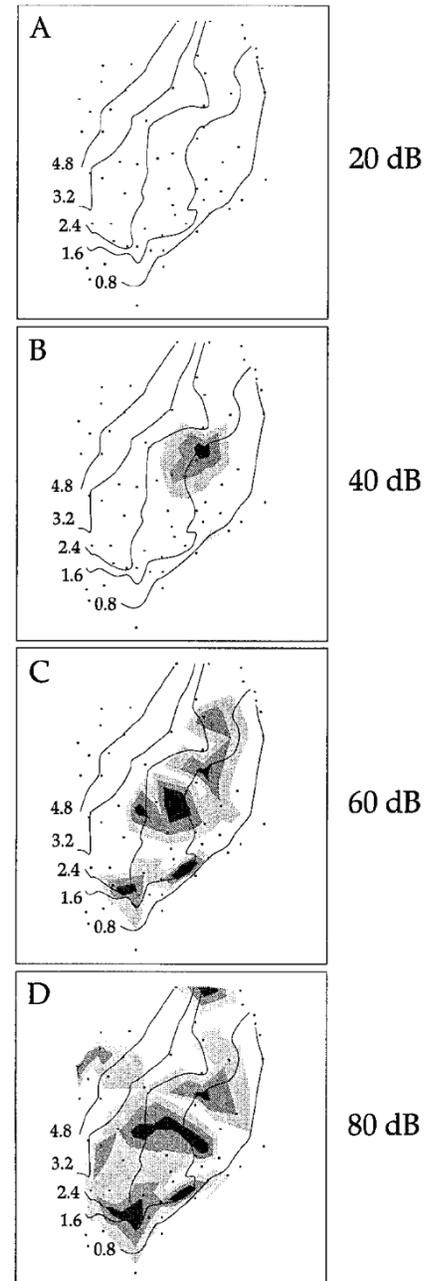
# Effects of Sound Level on Tonotopy



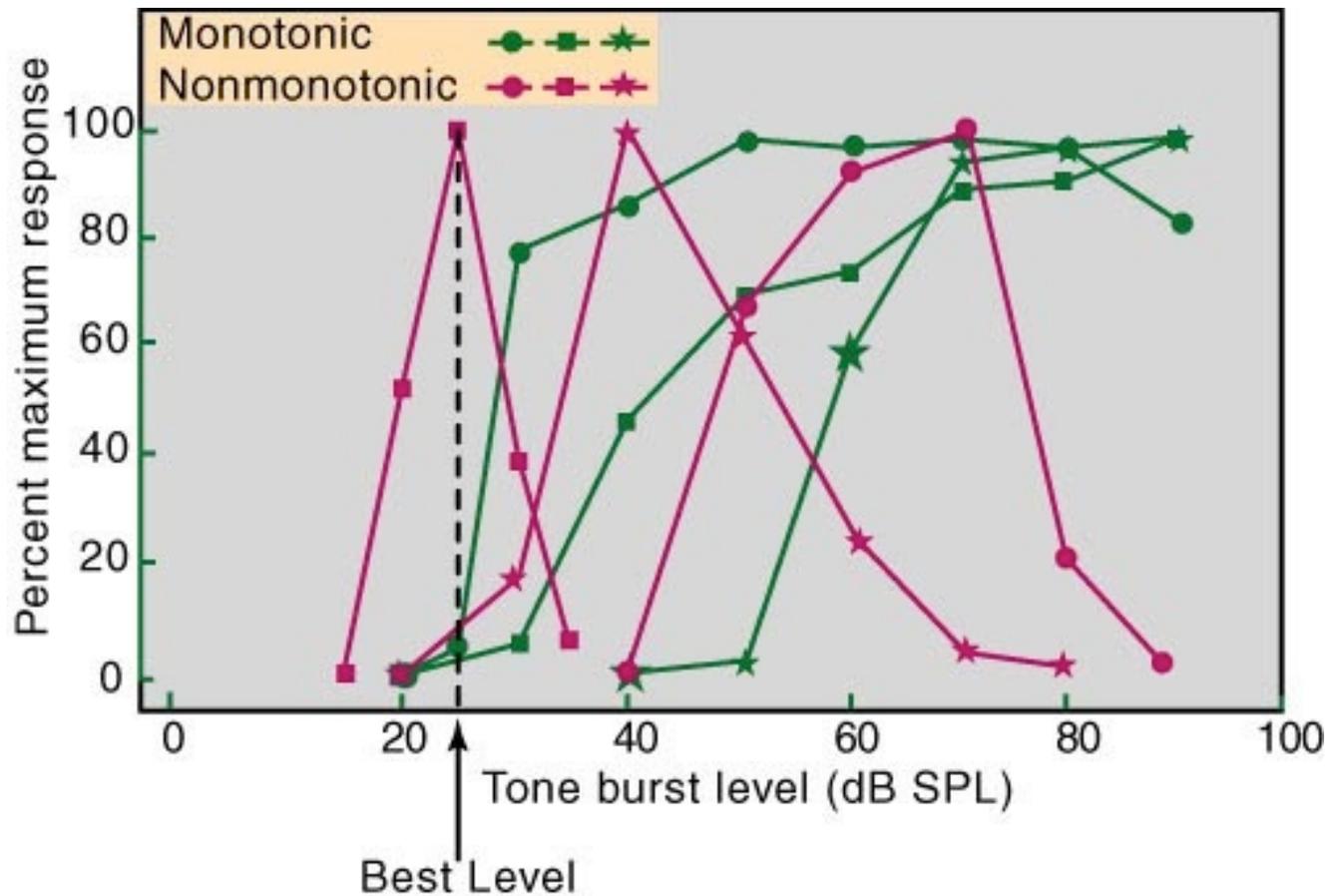
*Scattered distribution of response thresholds undermines iso-frequency representation*

02L001  
1.6 kHz

1.0 mm

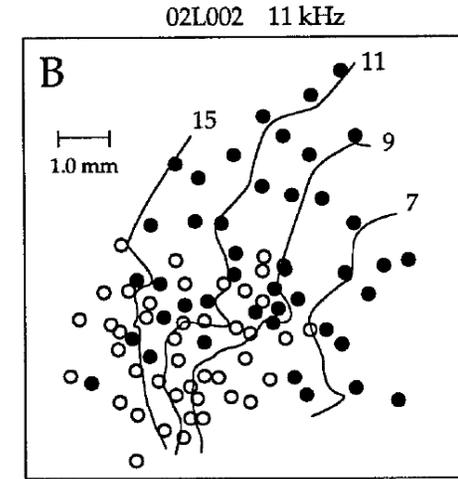
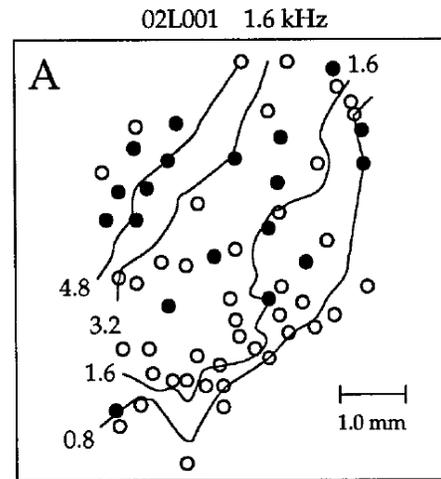
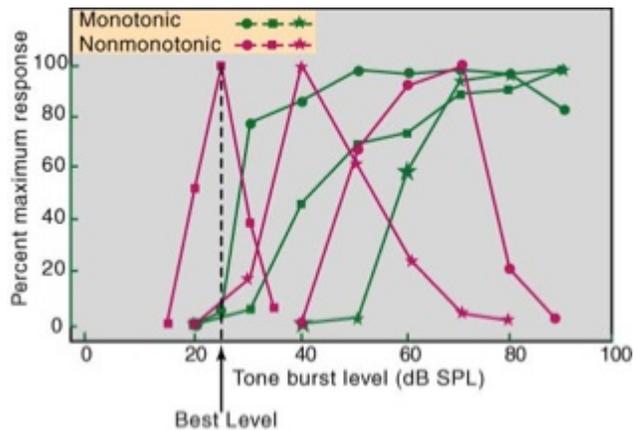


# Discharge Rate as a function of Sound Level: *Monotonic or Nonmonotonic*

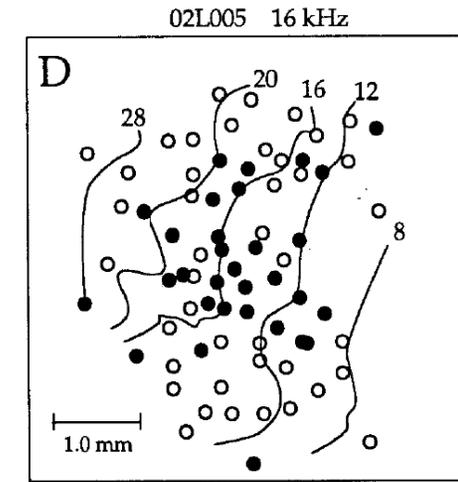
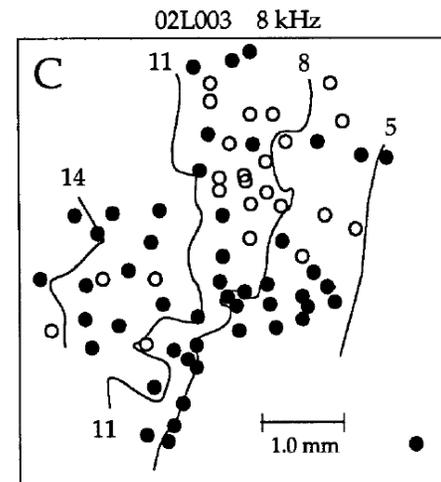


# Cortical Mapping of Sound Level

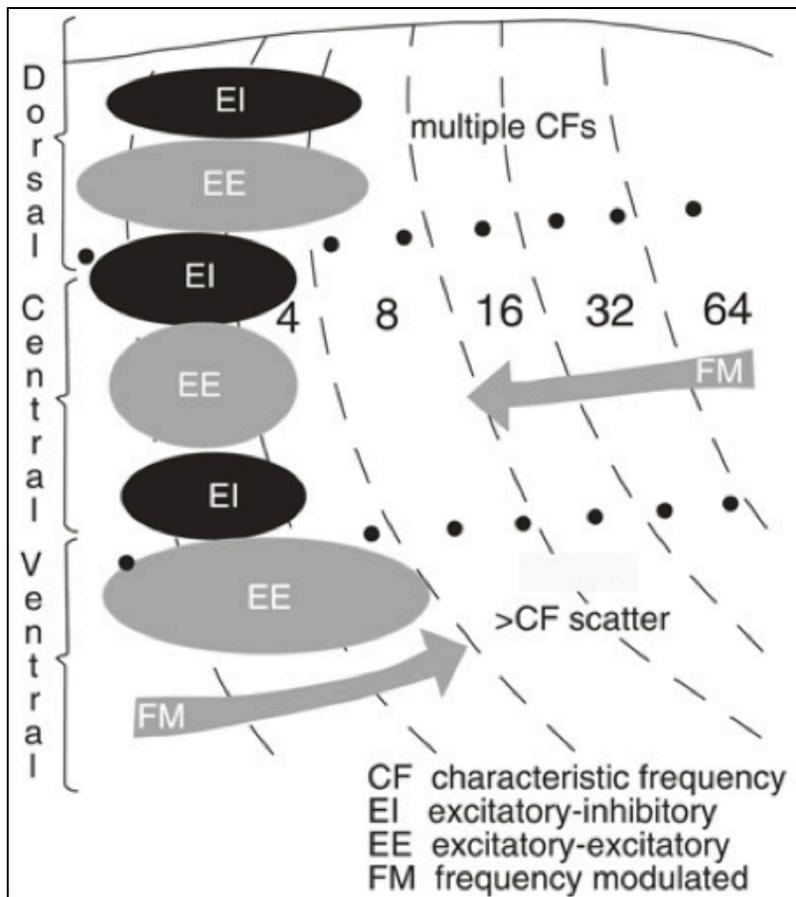
*clustering of cells with monotonic and nonmonotonic sensitivity to SPL*



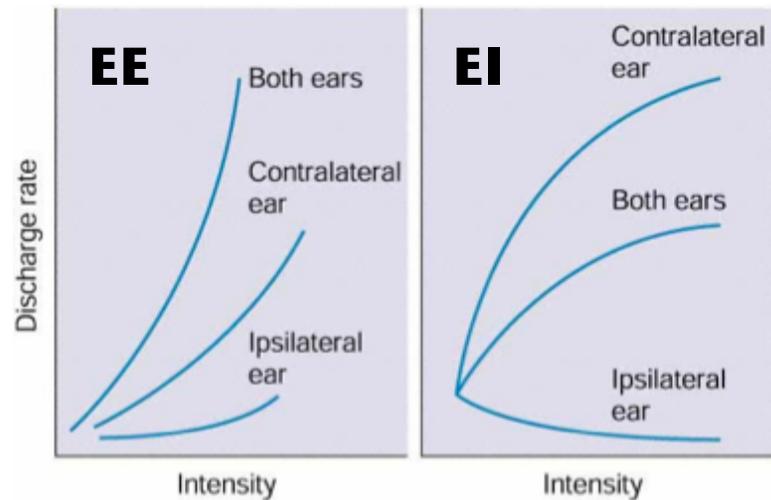
○ MONOTONIC  
● NONMONOTONIC



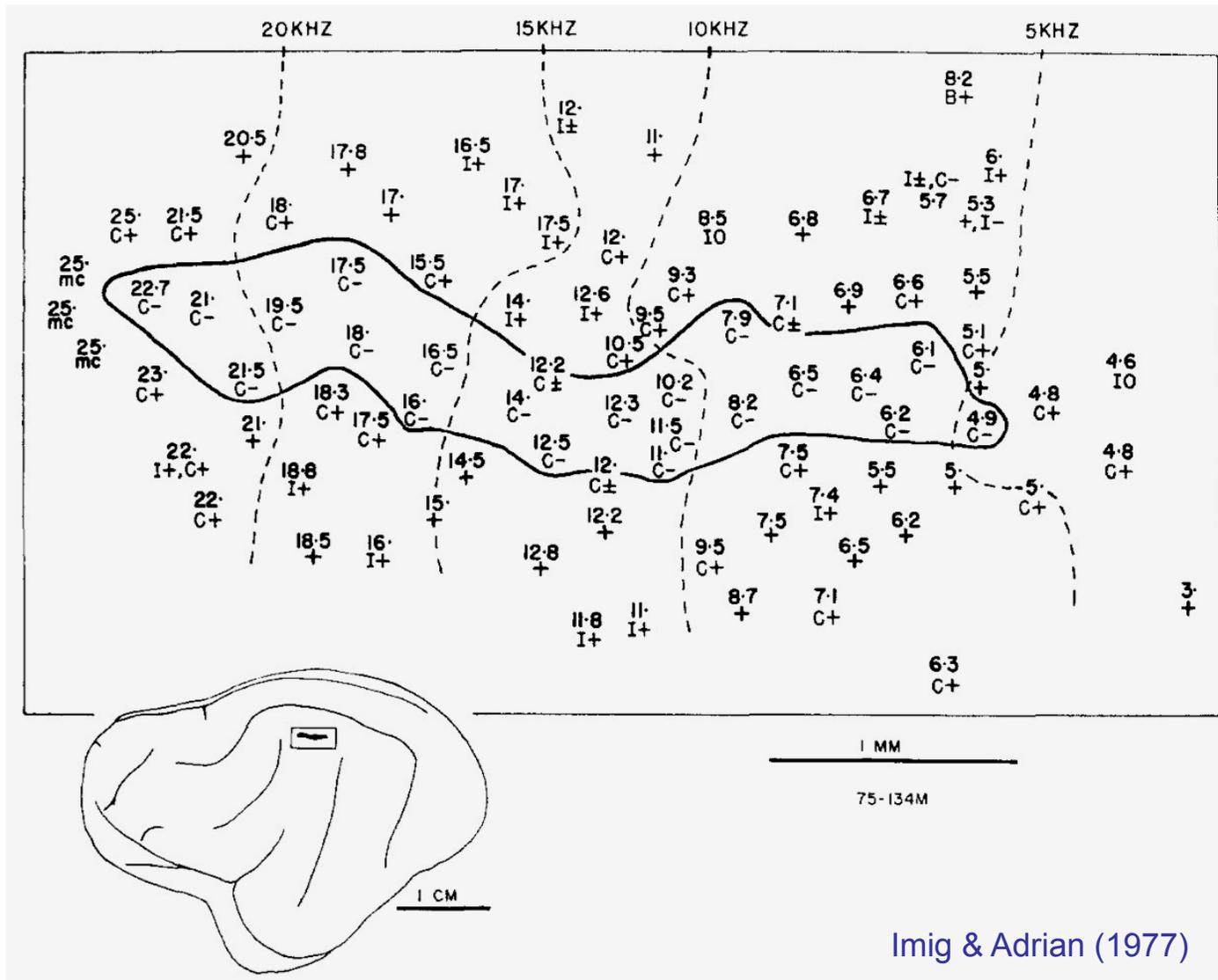
# Iso-frequency Functional Subregions of AI



Schreiner & Weiner (2007)

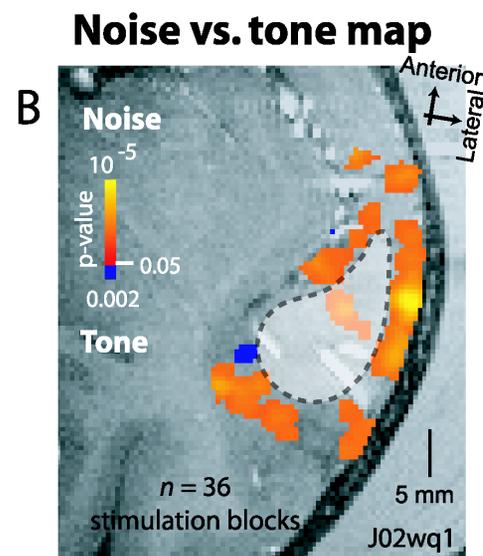
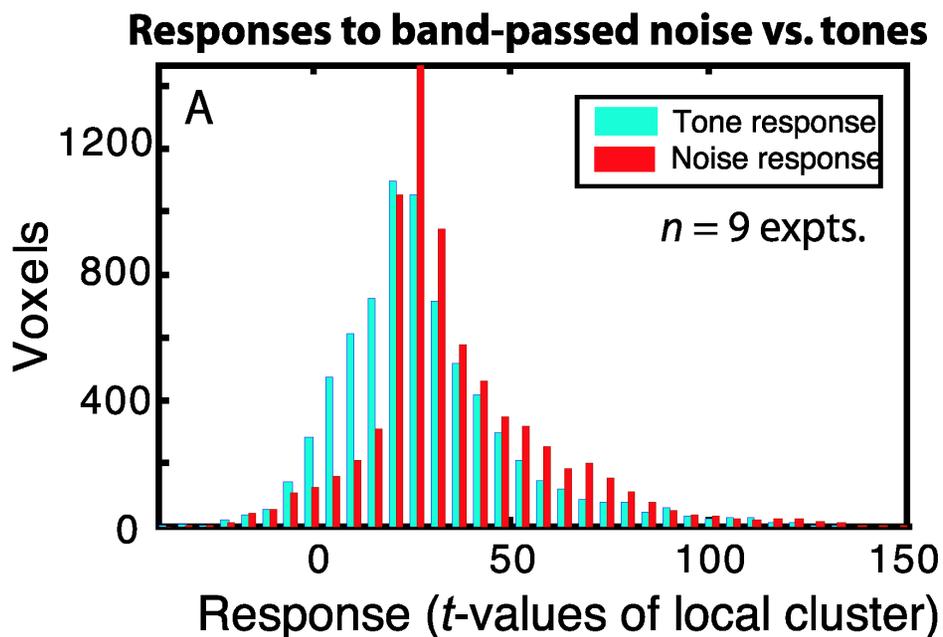
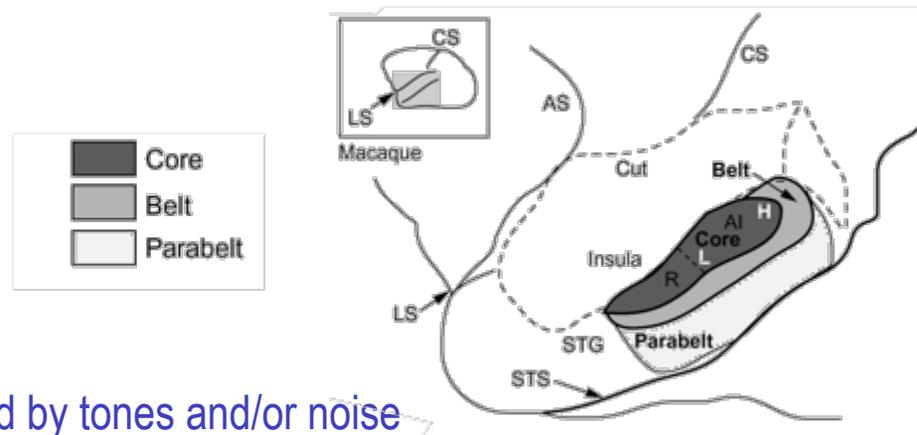


# Binaural Bands in cat AI



# Preference for Noise in Belt Auditory Cortex

- Analysis based on voxels significantly activated by tones and/or noise
- Voxels responded more to noise than tones
- Responses in auditory belt areas were larger to noise than tones



## Temporal, parietal, & frontal cortex

... which mediate space perception, auditory memory and other functions



## Parabelt

Rostral and caudal divisions lateral to the belt



## Belt

Multiple fields less precisely cochleotopic and more responsive to complex stimuli than tones.



## Core

Cochleotopically organized fields that generally display clear response to pure tones



## MGB Ventral nucleus

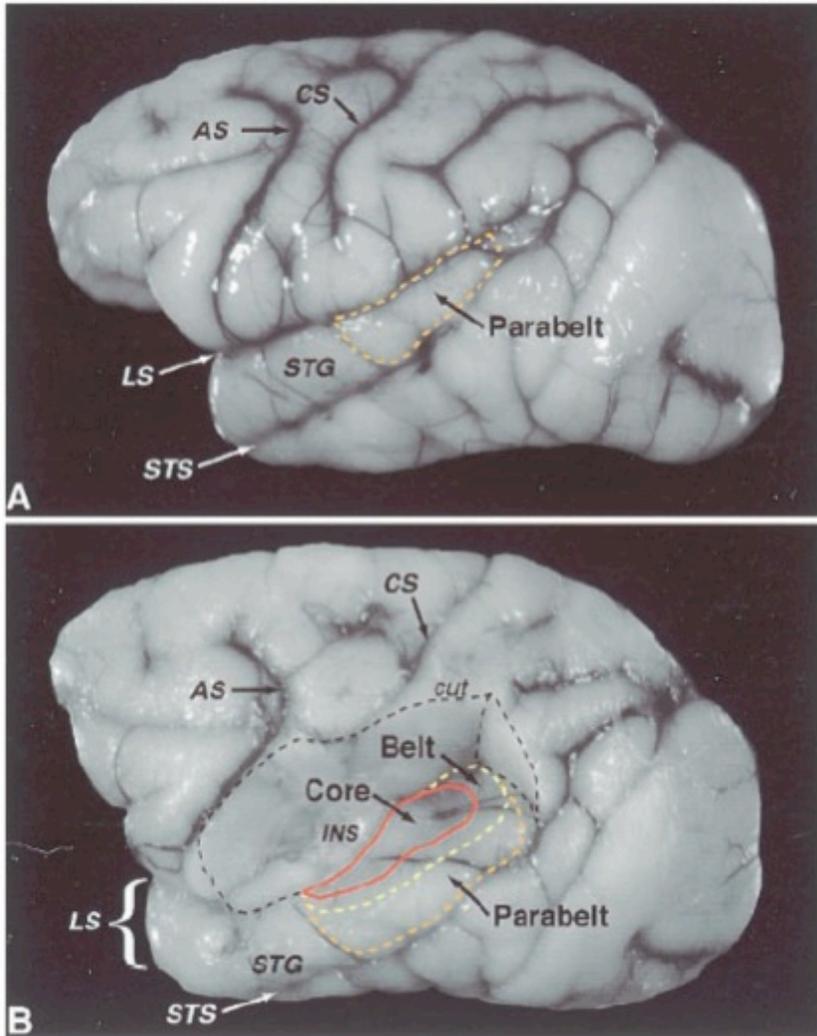
Principal division of the auditory thalamus

The belt and parabelt:

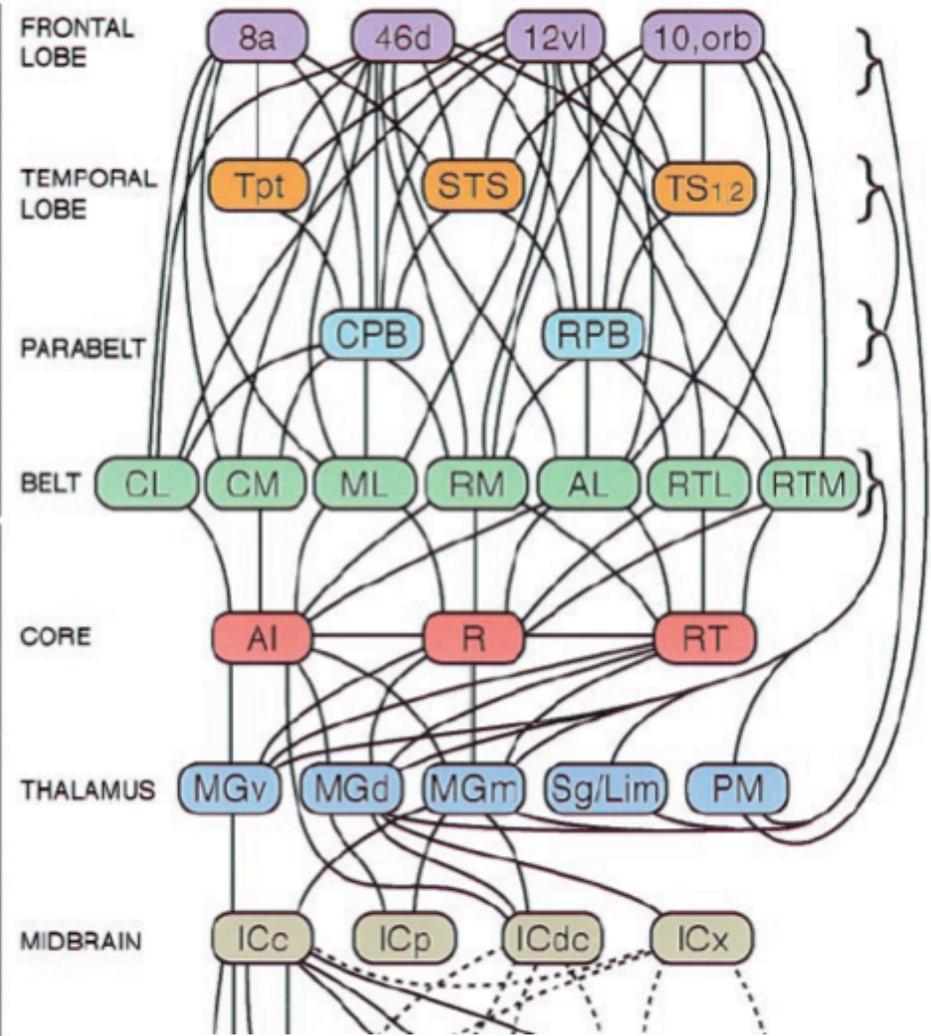
- have additional inputs from other MG nuclei and other parts of the thalamus.
- are implicated in integrative and associative functions involved in pattern perception and object recognition.

*Kaas, J.H., Hackett, T.A., and Tramo, M.J. Auditory processing in primate cerebral cortex, Current Opinion in Neurobiology, 1999, 9:164-170.*

# Auditory cortical subdivisions & processing streams



Macaque



Kaas & Hackett 2000

# Auditory Cortical Processing Streams

## **PATTERN** (“what”)

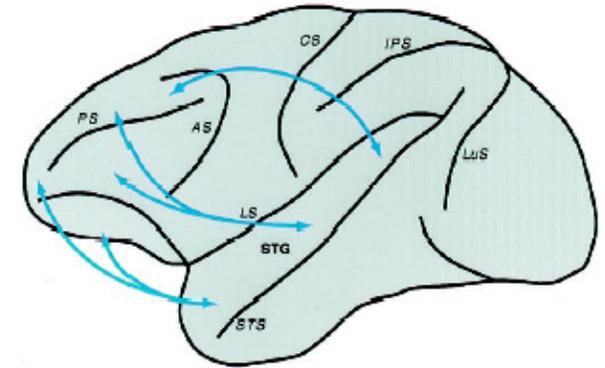
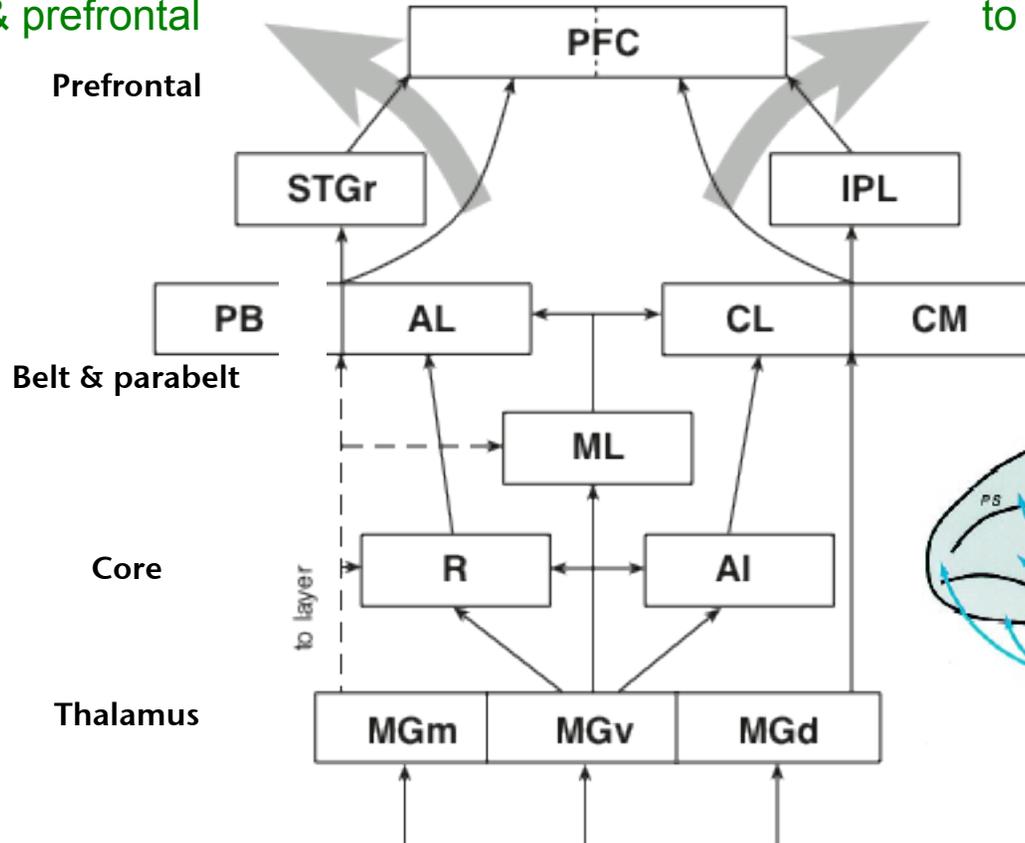
*thought to represent pattern/object information in a stream originating in anterior core and belt areas, broadly influencing the temporal lobe.*

rostral auditory cortex to temporal & prefrontal

## **SPACE** (“where”)

*thought to link AI via the caudomedial belt to frontal eye field and parietal targets implicated in spatial processing.*

caudal auditory cortex to parietal & prefrontal



# Auditory Cortical Processing Streams: Evidence?

***Evidence from human imaging studies generally supports the dual-pathway (streams) hypothesis...***

“What” processing (*e.g., determining the pitch, using working memory, phonemes*) predominantly activates auditory cortex and inferior prefrontal lobe.

Spatial processing is associated with the activation of posterior temporal areas, parietal cortex and superior frontal sulcus.

***...But imaging studies do reveal some caveats***

Zatorre et al (2001, 2002) showed that the ‘where’ pathway might be involved in tasks like sensorimotor integration and the disambiguation of overlapping sound sources, and that spatial sensitivity is linked to spectro-temporal features.

Alain et al. (2001) reported a significant correlation between the temporal and parietal cortex during pitch and localization tasks, suggesting considerable cross-talk between the streams.

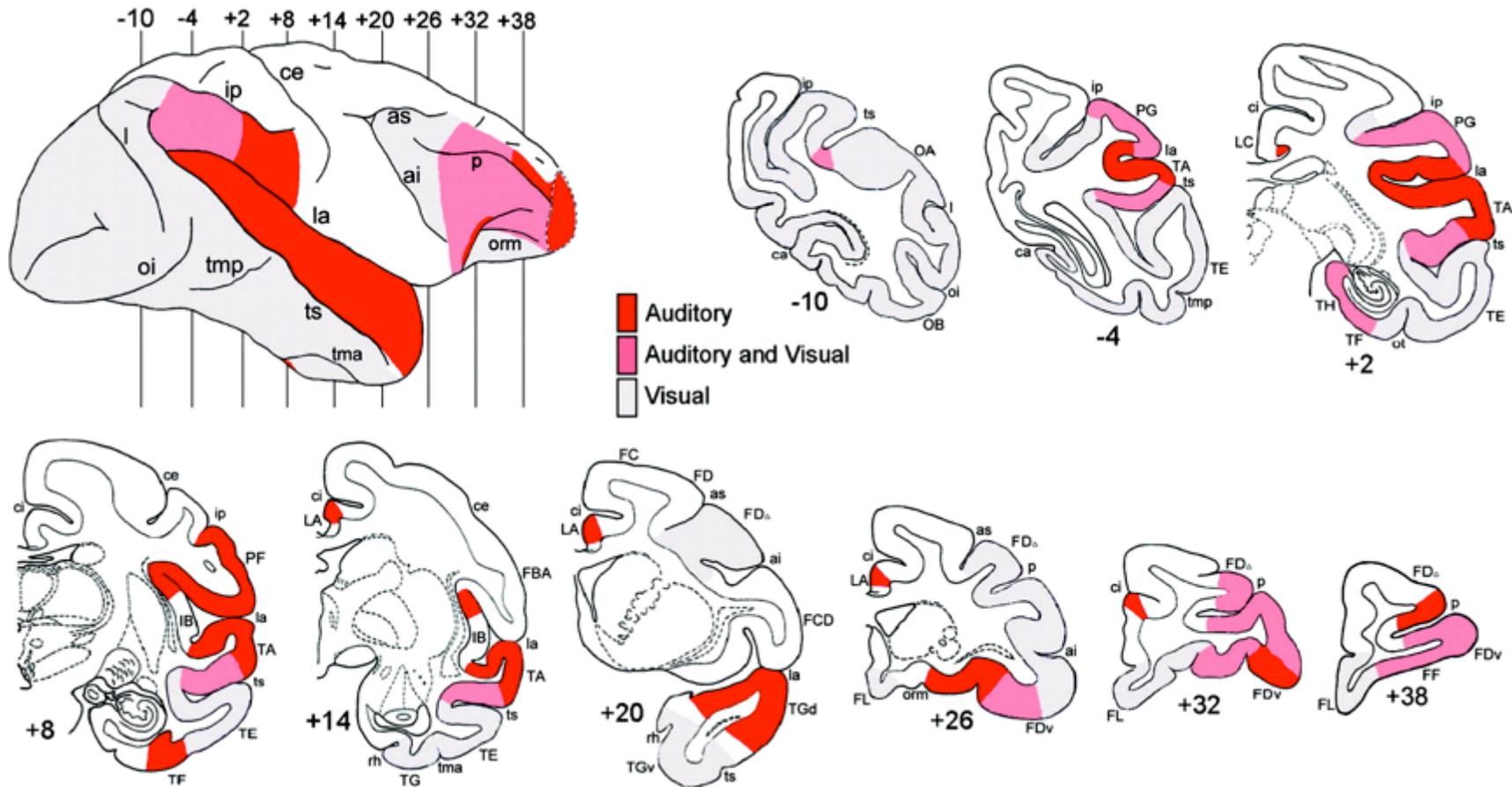
# Auditory Cortical Processing Streams: Evidence?

***Localization and pattern processing are not necessarily independent auditory tasks***

Spatial information is a key factor contributing to the “determination” (or recognition) of sound sources, and sound localization is greatly influenced by the spectral and temporal structure of sounds. Moreover, a growing body of evidence suggests that spatial information is carried in temporal discharge patterns and encoded in the ensemble activity distributed throughout auditory cortex.

# Multimodal Role for Dorsal Pathway?

*Poremba (2003) – functional mapping*



Schematic summary of cortical areas related to the processing of auditory, auditory plus visual, and visual stimuli. Based on 2-DG studies of unilaterally deaf macaques. (Poremba et al. 2003)

# Neuronal Mechanisms Underlying the Perception of Pitch and Harmony

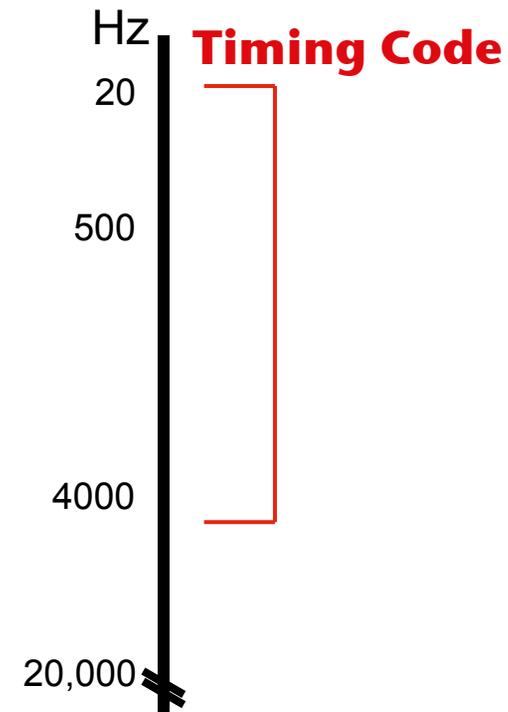
Ann. N.Y. Acad. Sci. 1060: 50–52 (2005)

GERALD LANGNER

**ABSTRACT:** Temporal processing of periodic acoustic signals in the auditory brain stem provides an explanation for pitch perception and the natural preference of our hearing system for harmonic relationships in music. Experimental evidence is reviewed for a corresponding neuronal model of correlation analysis and the spatial representation of pitch information along the second neural axis of the auditory system.

## Timing theories:

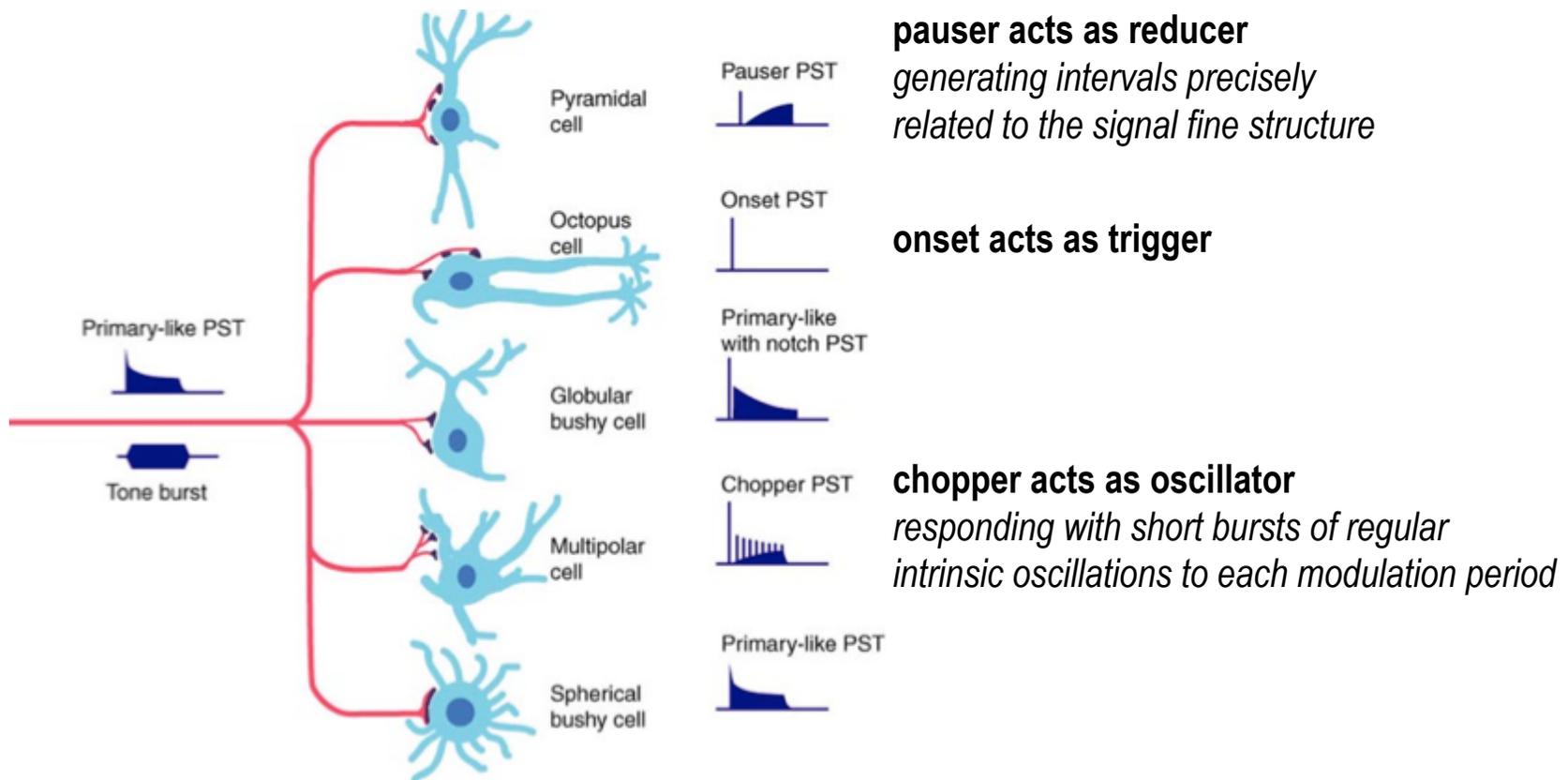
- sufficient for musical pitch, which deteriorates at high frequencies



# Langner's view of periodicity pitch processing

*a correlation is performed between signal fine-structure and envelope*

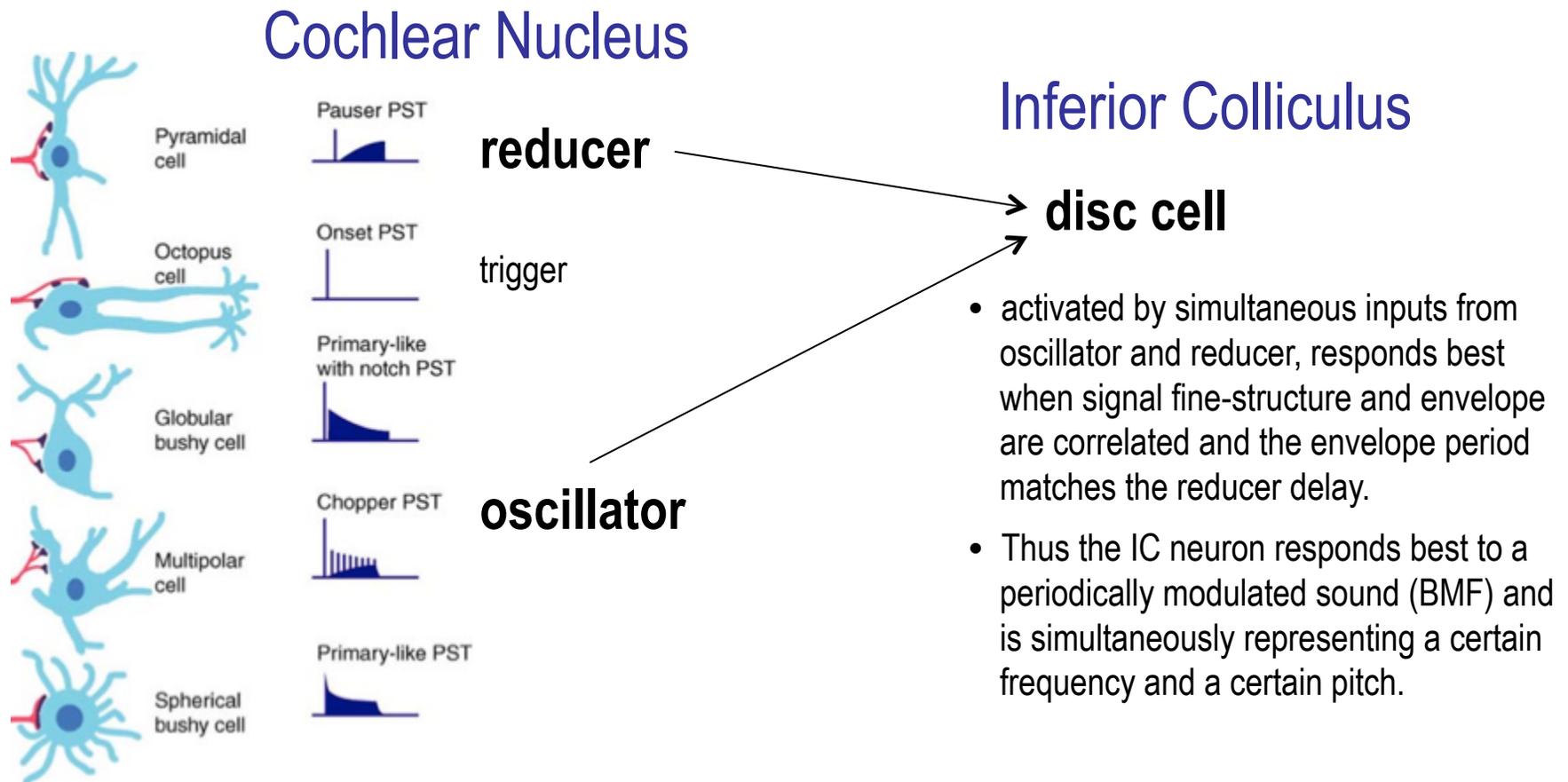
Within Cochlear Nucleus:



Langner, G. 1992. Periodicity coding in the auditory system. *Hearing Res.* 60: 115–142.

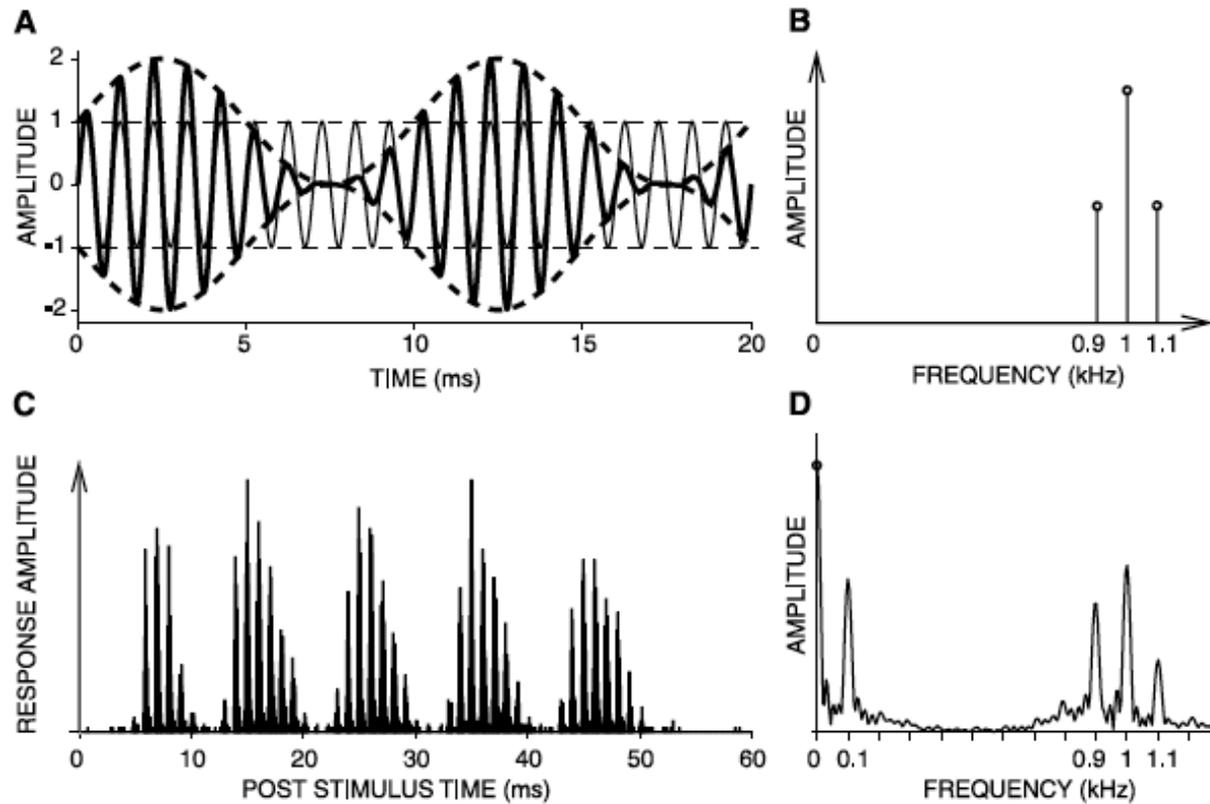
# Langner's view of periodicity pitch processing

*a correlation is performed between signal fine-structure and envelope*



# Neural Representation of AM

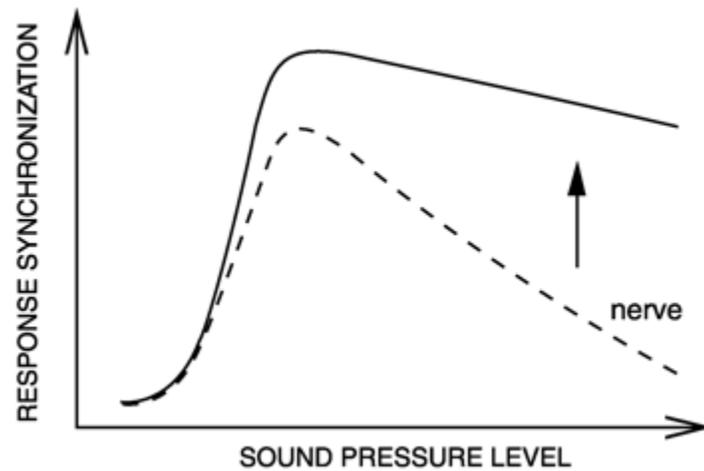
*auditory nerve*



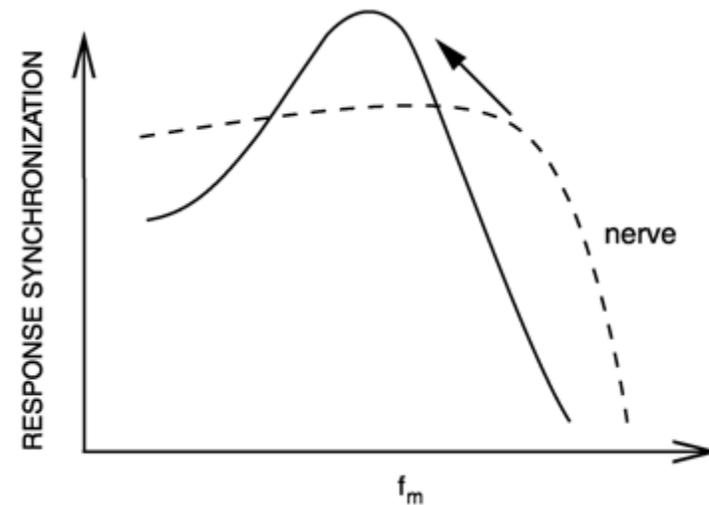
# Transformation in Representation of AM

*auditory nerve (dashed lines) and cochlear nucleus (solid lines)*

*enhancement of envelope synchronization and extended dynamic range*

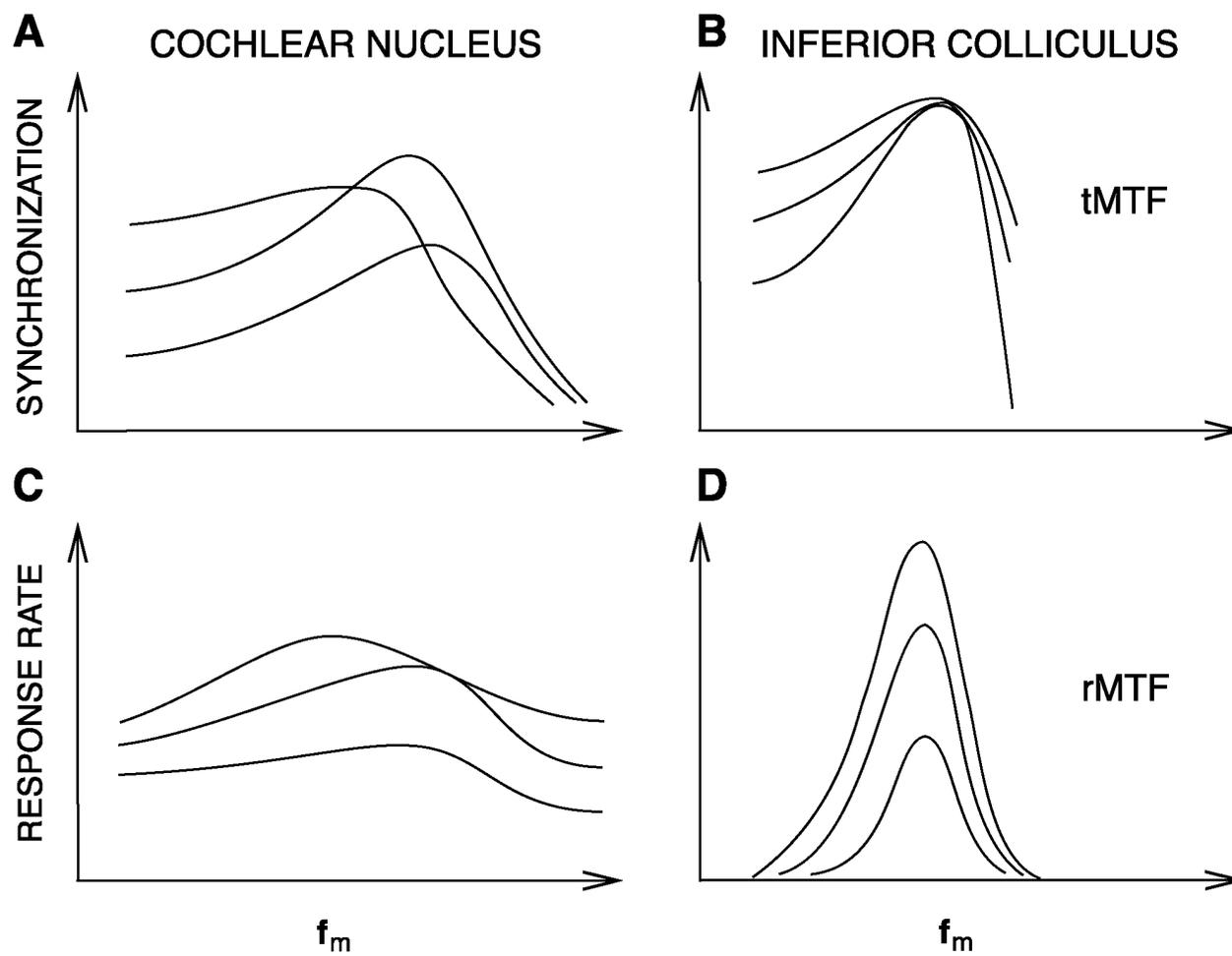


*Emergence of bandpass tMTFs.*



# Transformation in representation of AM

*Brainstem to midbrain*



# Auditory Midbrain Rate Tuning to AM

Langner and Schreiner (1988)

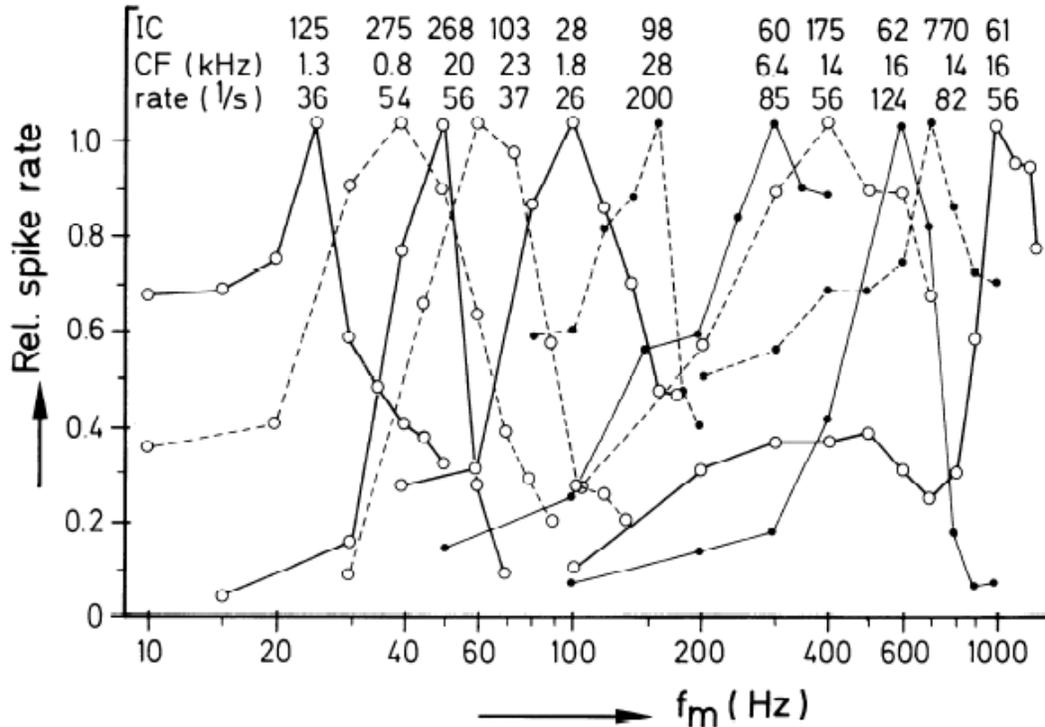


FIG. 3. Normalized modulation transfer functions (MTFs) for the firing rate of typical bandpass units of the ICC. Numbers above the MTFs indicate the unit number; the CF of the unit that also serves as carrier frequency of the AM stimuli; and the maximal firing rate corresponding to the tip of the MTF. Single units are indicated by open circles, multiple units by closed circles.

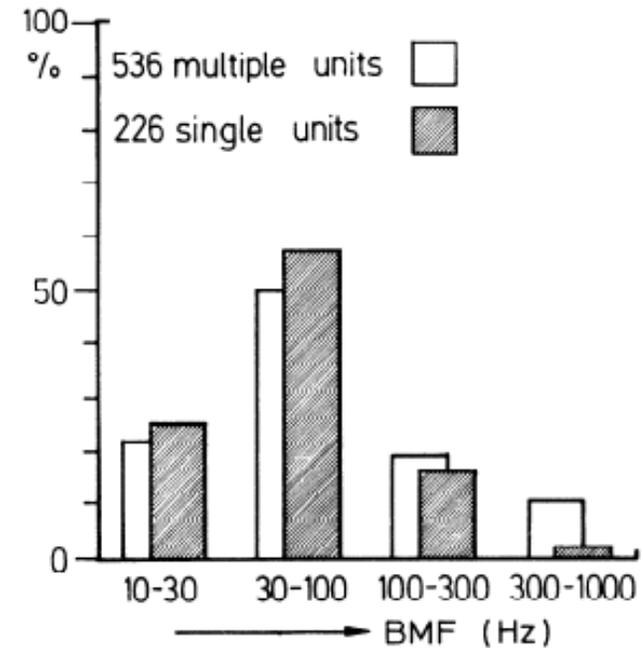
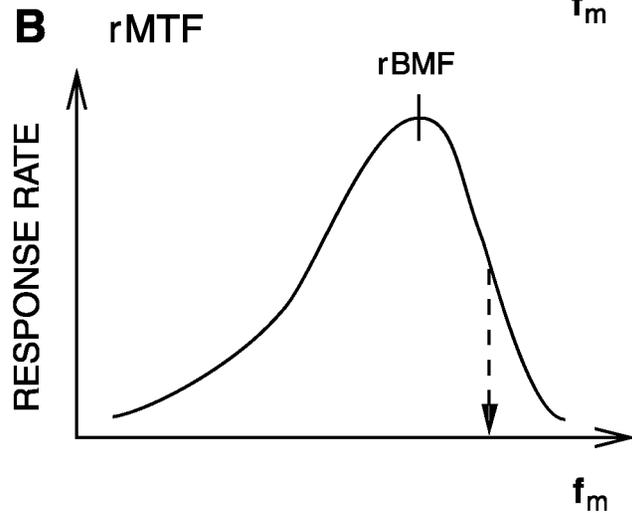
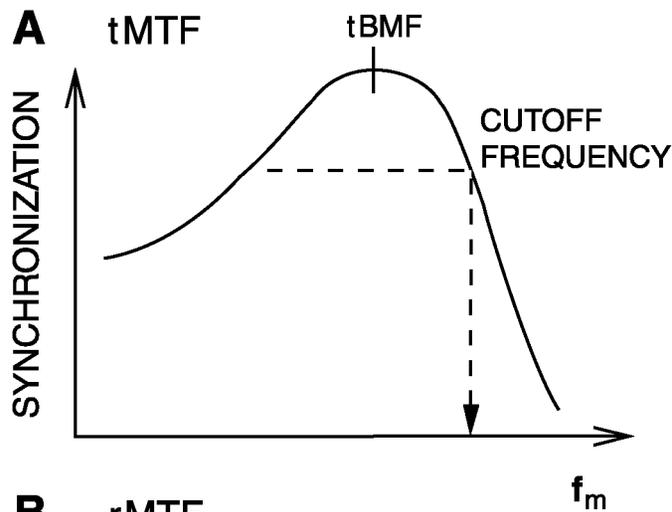


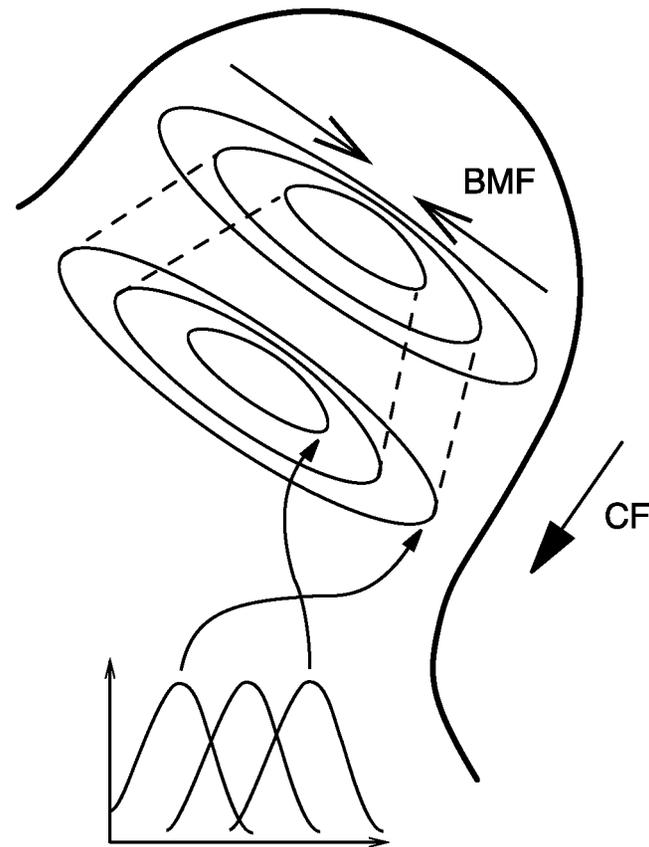
FIG. 1. Distribution of best modulation frequencies (rBMFs) for 536 multiple units and 226 single units in the inferior colliculus. About one-half of the recordings revealed BMFs between 30 and 100 Hz.

# Topographic Representation of AM

## INFERIOR COLLICULUS

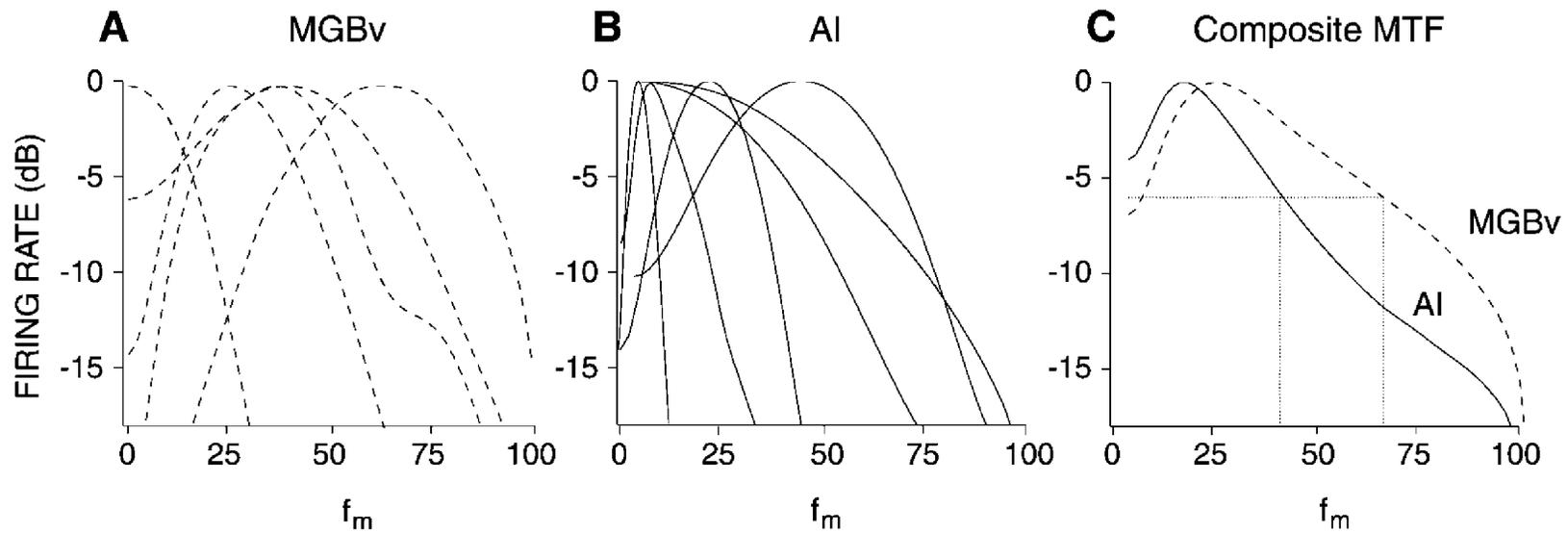


## **C** BMF TOPOGRAPHY



# Transformation in representation of AM

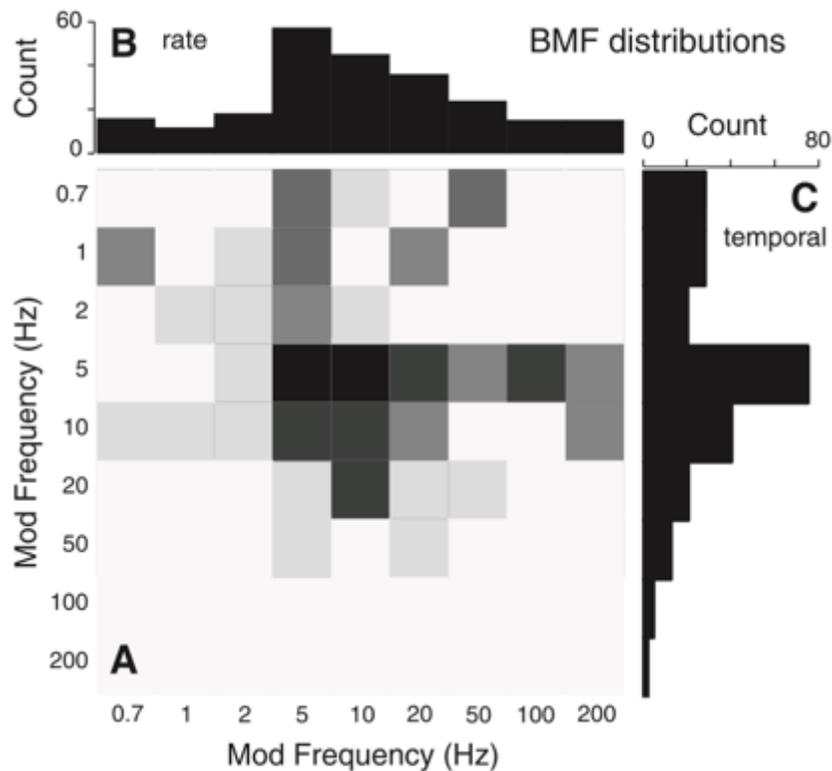
## *Thalamus to cortex*



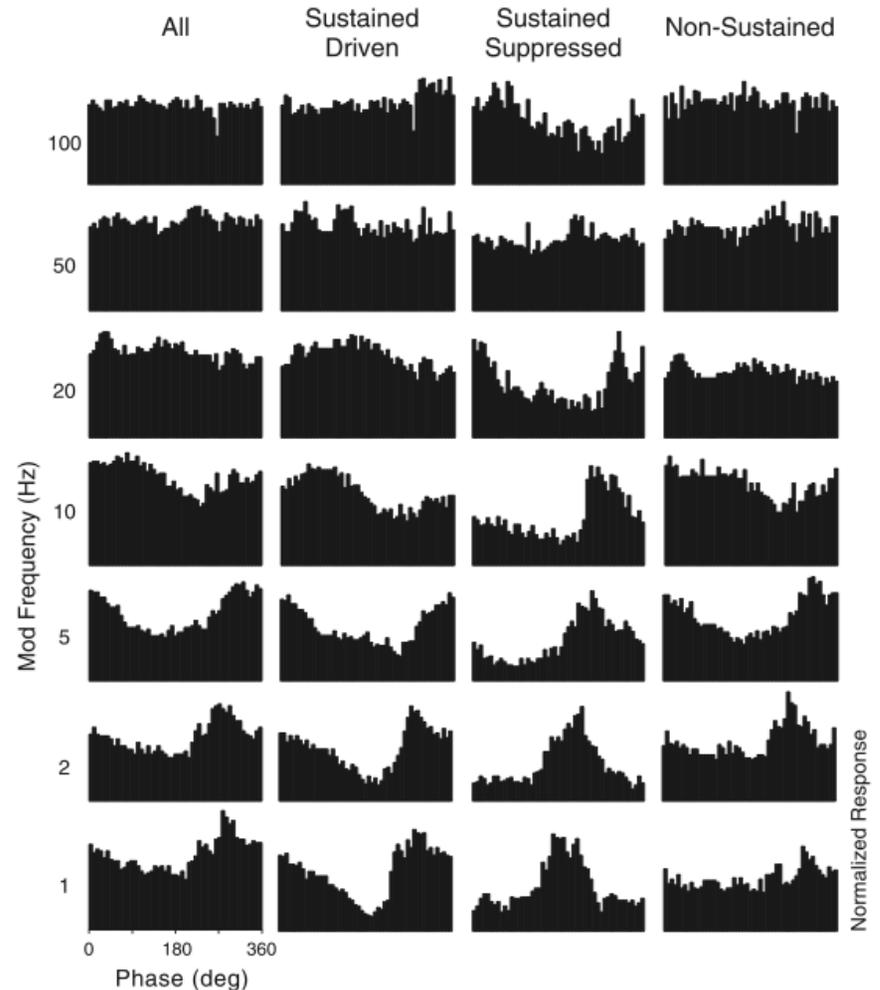
Typical tMTFs (synchronized firing rate) from MGB and AI (anesthetized cat).

Composite tMTFs derive from averaging all tMTFs for thalamic and cortical units separately. The dotted lines indicate the 6-dB upper cut-off frequency.

# AM Coding in Rhesus Auditory Cortex

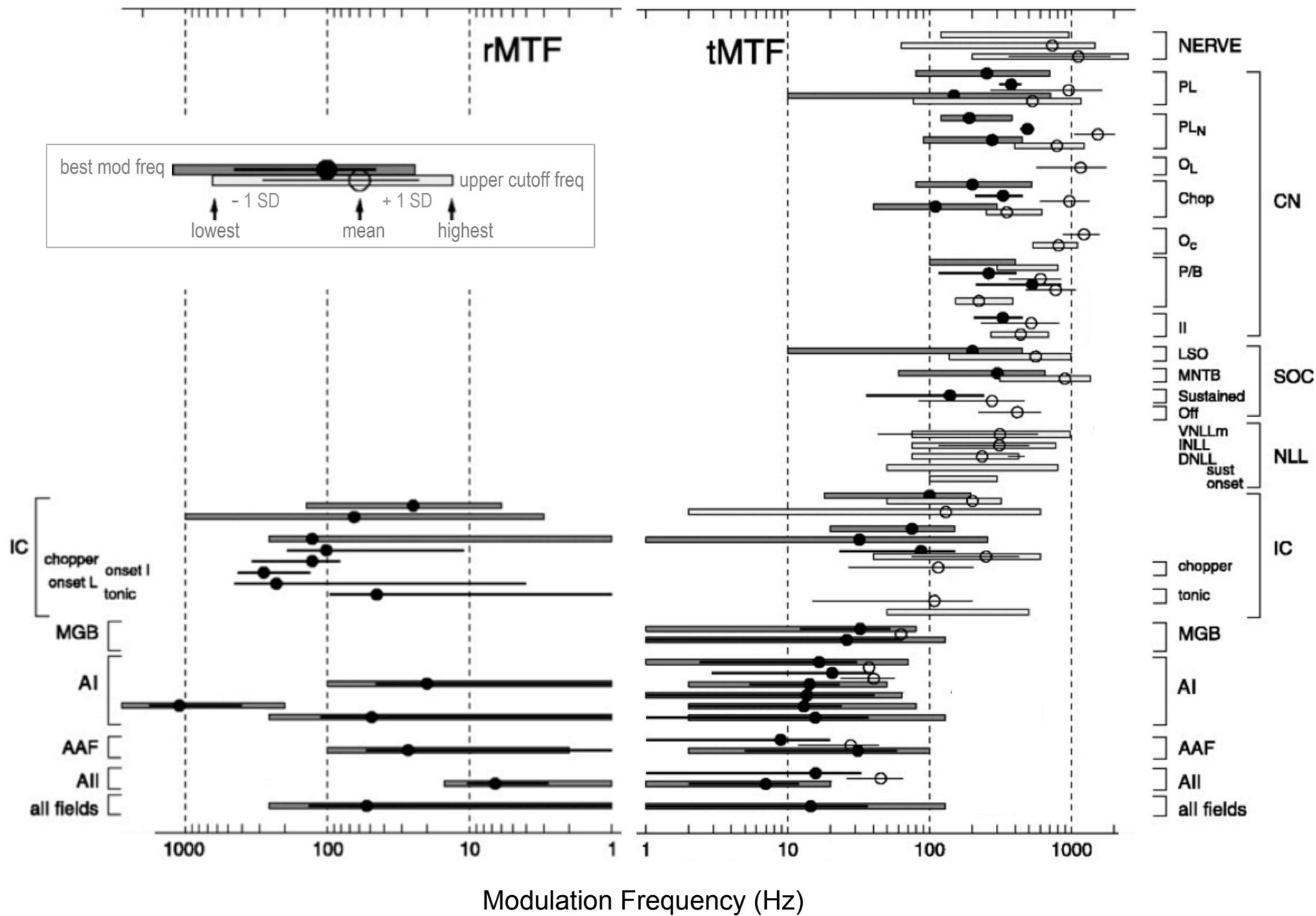


Macaque auditory cortical neurons respond with highest firing rate and best synchrony to slowly modulating SAM stimuli (<20 Hz)



Composite MPHs show how activity of a population of neurons recorded from the same hemisphere of one rhesus monkey responds to 60-dB SAM stimuli at best carrier frequency.

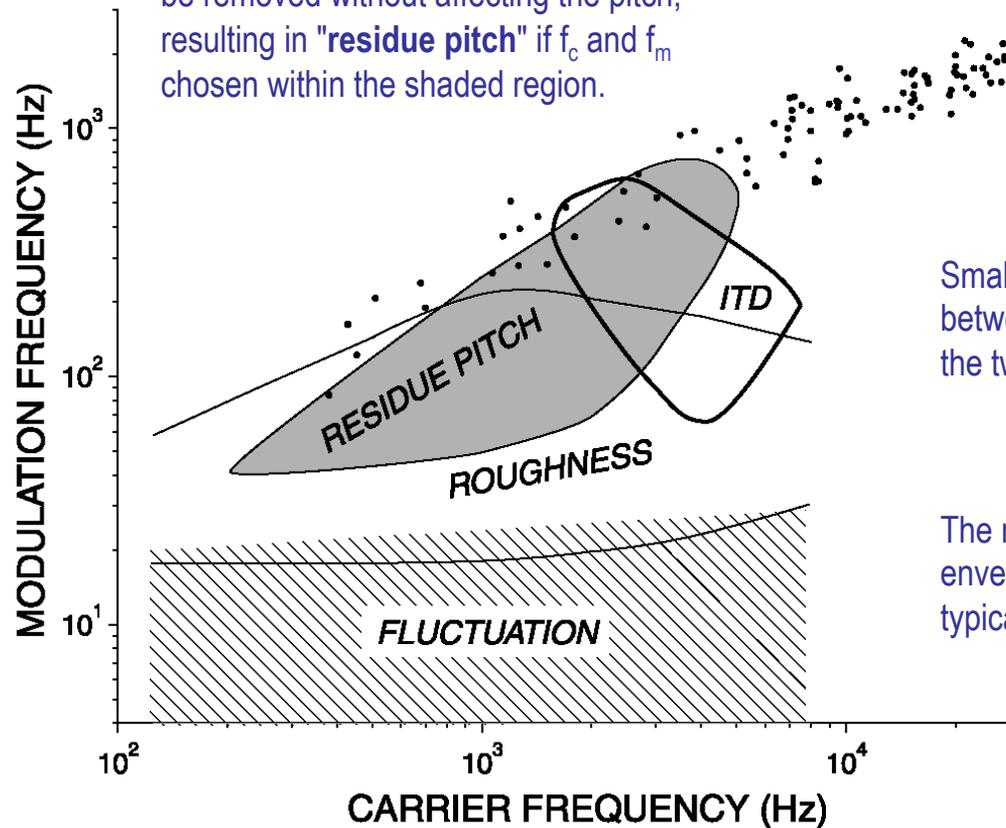
Malone et al (2007)



# Percepts of Amplitude-Modulated Tones

Harmonic complex tones produce a pitch comparable to the fundamental frequency, but the lower harmonics can be removed without affecting the pitch, resulting in "residue pitch" if  $f_c$  and  $f_m$  chosen within the shaded region.

Dots indicate cutoff values for modulation transfer functions of auditory nerve fibers



Small ITDs can be detected between modulated stimuli to the two ears.

The rate at which the temporal envelope of fluent speech varies is typically about 4 Hz (syllables/s)

---

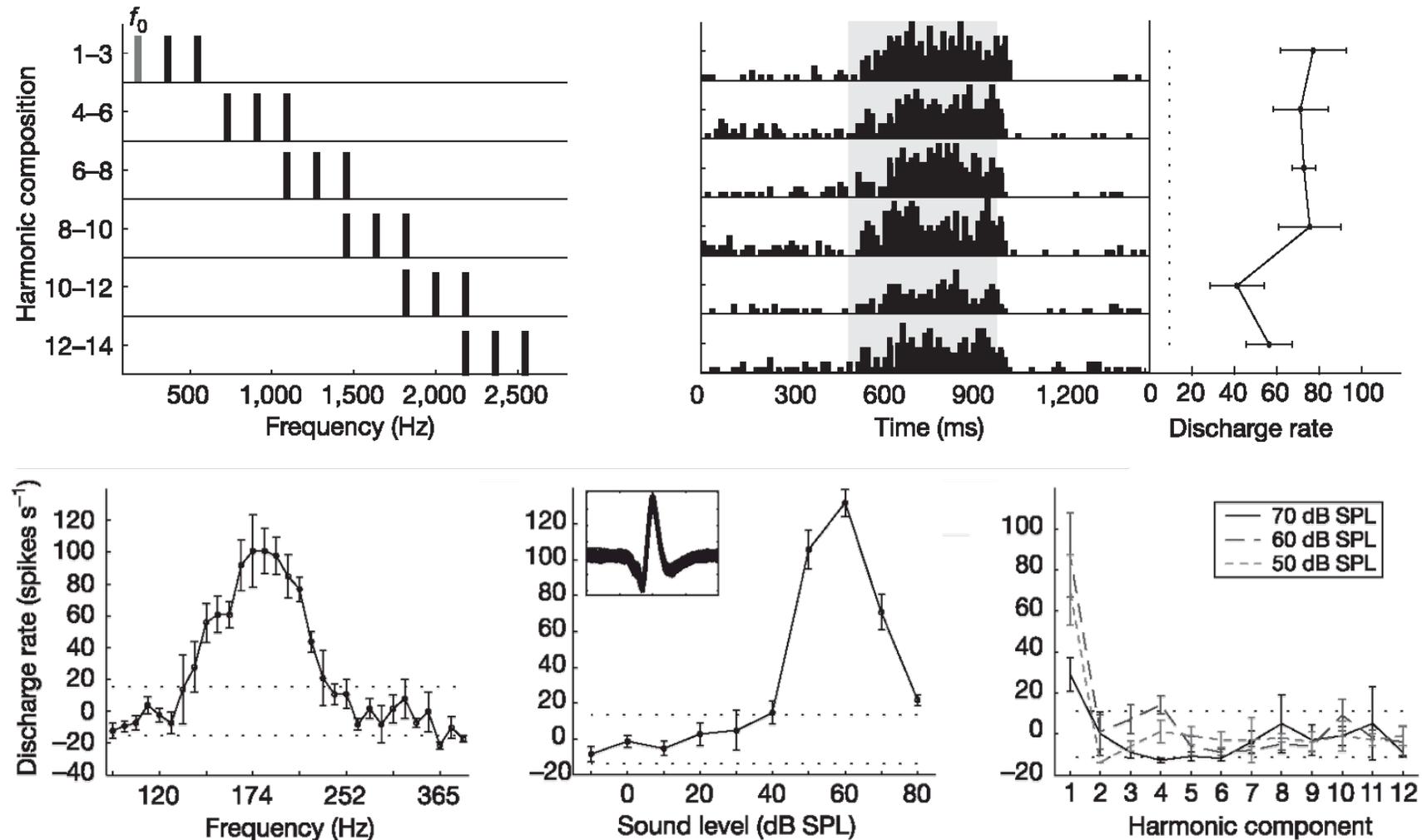
# The neuronal representation of pitch in primate auditory cortex

Daniel Bendor<sup>1</sup> & Xiaoqin Wang<sup>1</sup>

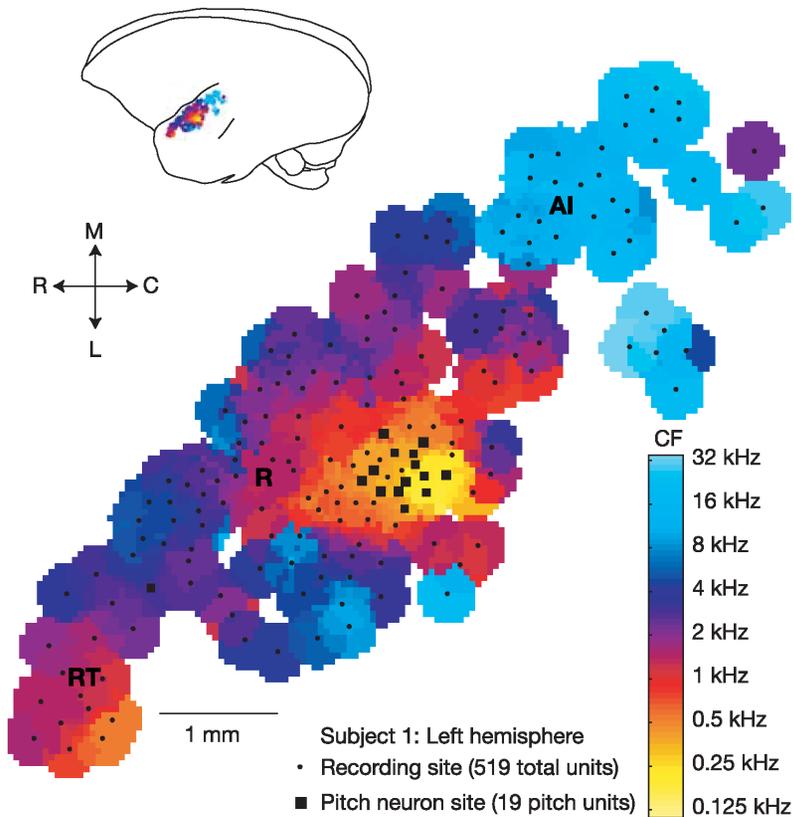
Pitch perception is critical for identifying and segregating auditory objects<sup>1</sup>, especially in the context of music and speech. The perception of pitch is not unique to humans and has been experimentally demonstrated in several animal species<sup>2,3</sup>. Pitch is the subjective attribute of a sound's fundamental frequency ( $f_0$ ) that is determined by both the temporal regularity and average repetition rate of its acoustic waveform. Spectrally dissimilar sounds can have the same pitch if they share a common  $f_0$ . Even when the acoustic energy at  $f_0$  is removed ('missing fundamental') the same pitch is still perceived<sup>1</sup>. Despite its importance for hearing, how pitch is represented in the cerebral cortex is unknown. Here we show the existence of neurons in the auditory cortex of marmoset monkeys that respond to both pure tones and missing fundamental harmonic complex sounds with the same  $f_0$ , providing a neural correlate for pitch constancy<sup>1</sup>. These pitch-selective neurons are located in a restricted low-frequency cortical region near the anterolateral border of the primary auditory cortex, and is consistent with the location of a pitch-selective area identified in recent imaging studies in humans<sup>4,5</sup>.

**Figure 1 | An example of a pitch-selective neuron (unit M36n-532).** Error bars represent standard error of the mean (s.e.m.). The dotted black lines indicate the significance level for discharge rate ( $\pm 2$  standard deviations away from the spontaneous discharge rate). **a**, Frequency spectra of a series of harmonic complex stimuli. The fundamental frequency component ( $f_0$ ) and its higher harmonics have equal amplitudes of 50 dB SPL. **b**, Peristimulus time histogram (left) and tuning curve (right) of the neuron's response to the stimuli in **a**. Stimuli were presented from 500 to 1,000 ms (indicated by the shaded region on the left plot). **c**, Frequency

tuning of the neuron derived from pure tones. **d**, Response of the neuron to a pure tone at characteristic frequency (182 Hz) across sound levels (rate-level function). Inset plot shows an overlay of 2,434 digitized action potentials recorded from this neuron (displayed within a 2 ms window). **e**, The neuron's responses to individual harmonics (number 1–12) at three sound levels, respectively. All the harmonics above the  $f_0$  component (first harmonic) were outside the neuron's excitatory frequency response area, and did not elicit significant responses. SPL, sound pressure level.

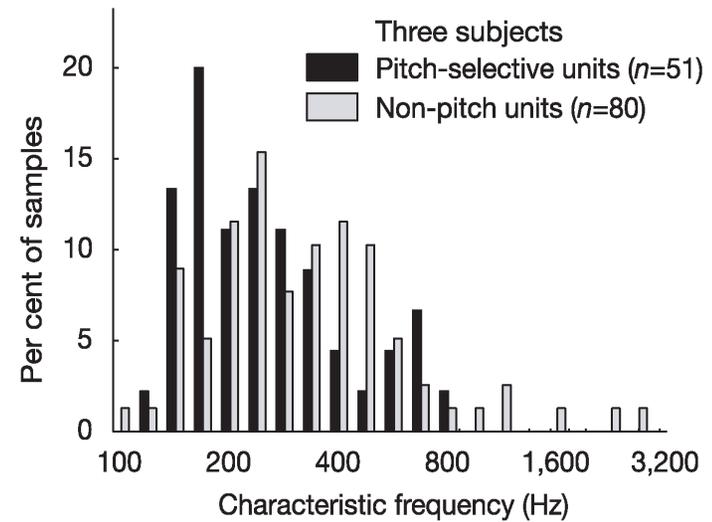


53 neurons (3 marmosets) were “pitch-selective”.  
51/53 were located within a low-frequency region  
near the rostralateral border of AI.



*In this pitch region:*

- 131 neurons responded to pure tones, and 51 (39%) were pitch-selective. Pitch-selective and -nonselective neurons exhibited similar CFs.



- 50 additional neurons responded poorly to pure tones, but did respond to narrowband or wideband stimuli such as harmonic complex tones, SAM, SFM, click trains, or band-pass noise.
- A subset of 10 neurons only responded to harmonic complex and SAM tones with rep rates similar in frequency to the CFs of neighboring neurons.

# Binary Spiking

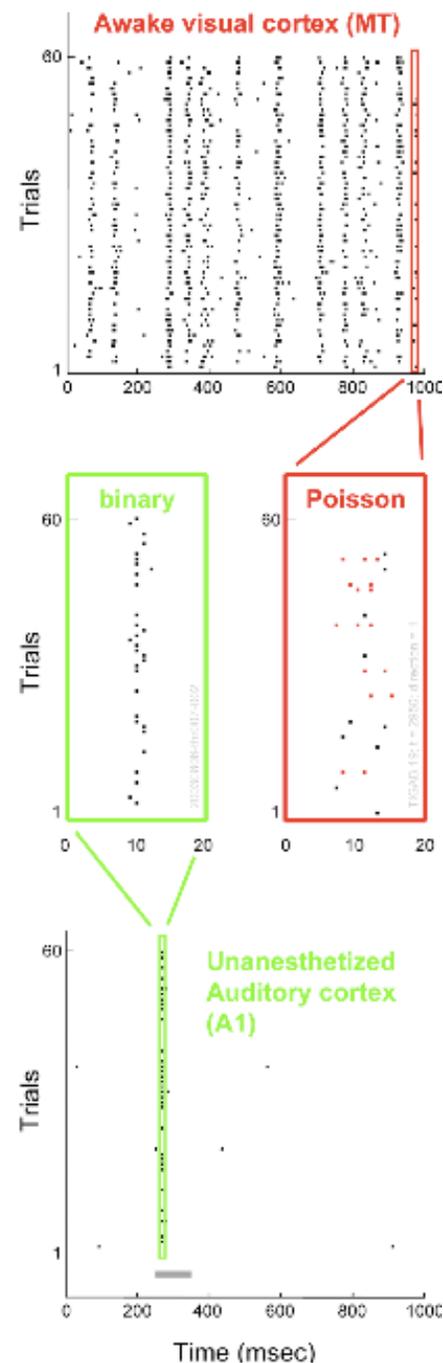
Neuron, Vol. 48, 479–488, November 3, 2005, Copyright ©2005 by Elsevier Inc.

## Reliability and Representational Bandwidth in the Auditory Cortex

Michael R. DeWeese, Tomáš Hromádka,  
and Anthony M. Zador\*

Cold Spring Harbor Laboratory  
Watson School of Biological Sciences  
Cold Spring Harbor, New York 11724

It is unclear why there are so many more neurons in sensory cortex than in the sensory periphery. One possibility is that these “extra” neurons are used to overcome cortical noise and faithfully represent the acoustic stimulus. Another possibility is that even after overcoming cortical noise, there is “excess representational bandwidth” available and that this bandwidth is used to represent conjunctions of auditory and nonauditory information for computation. Here, we discuss recent data about neuronal reliability in auditory cortex showing that cortical noise may not be as high as was previously believed. Although at present, the data suggest that auditory cortex neurons can be more reliable than those in the visual cortex, we speculate that the principles governing cortical computation are universal and that visual and other cortical areas can also exploit strategies based on similarly high-fidelity activity.

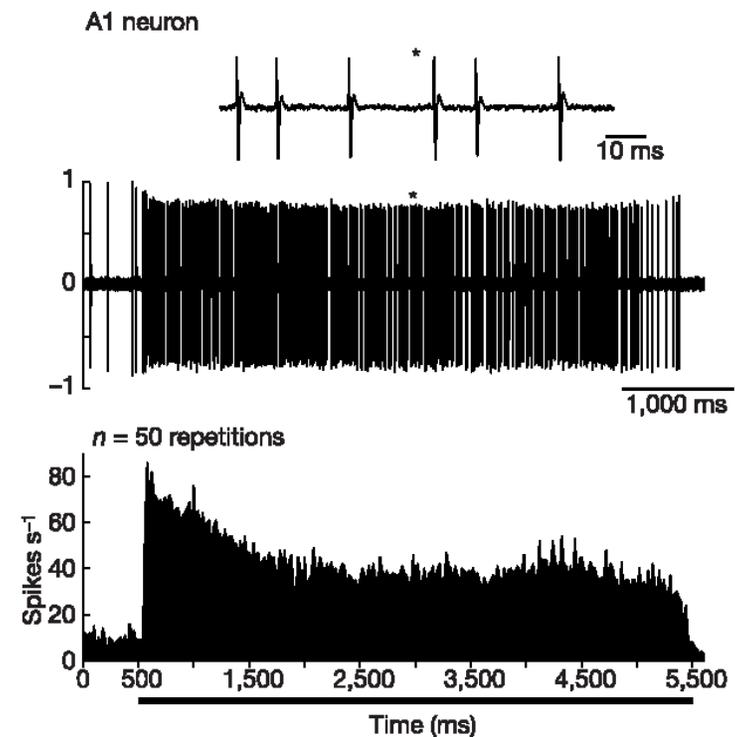


## LETTERS

# Sustained firing in auditory cortex evoked by preferred stimuli

Xiaoqin Wang<sup>1</sup>, Thomas Lu<sup>1</sup>, Ross K. Snider<sup>1†</sup> & Li Liang<sup>1</sup>

It has been well documented that neurons in the auditory cortex of anaesthetized animals generally display transient responses to acoustic stimulation, and typically respond to a brief stimulus with one or fewer action potentials<sup>1–5</sup>. The number of action potentials evoked by each stimulus usually does not increase with increasing stimulus duration<sup>1,5–7</sup>. Such observations have long puzzled researchers across disciplines and raised serious questions regarding the role of the auditory cortex in encoding ongoing acoustic signals. Contrary to these long-held views, here we show that single neurons in both primary (area A1) and lateral belt areas of the auditory cortex of awake marmoset monkeys (*Callithrix jacchus*) are capable of firing in a sustained manner over a prolonged period of time, especially when they are driven by their preferred stimuli. In contrast, responses become more transient or phasic when auditory cortex neurons respond to non-preferred stimuli. These findings suggest that when the auditory cortex is stimulated by a sound, a particular population of neurons fire maximally throughout the duration of the sound. Responses of other, less optimally driven neurons fade away quickly after stimulus onset. This results in a selective representation of the sound across both neuronal population and time.

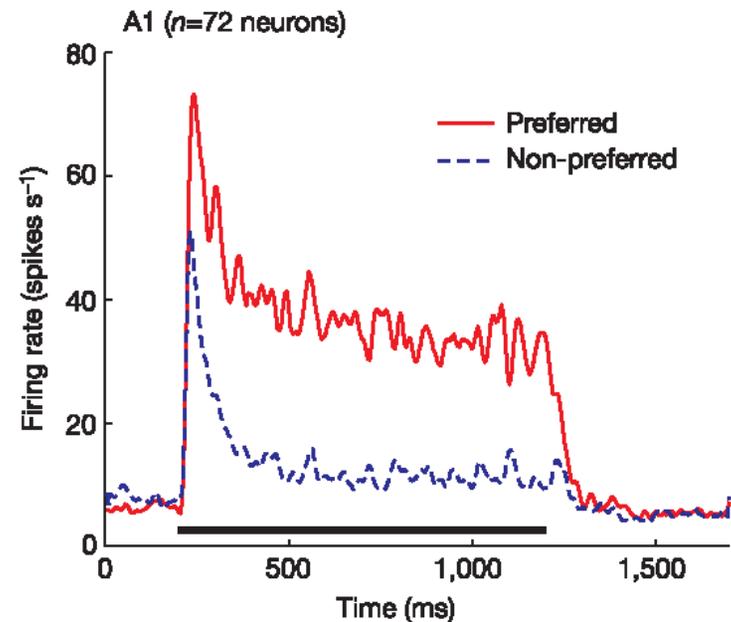


## LETTERS

## Sustained firing in auditory cortex evoked by preferred stimuli

Xiaoqin Wang<sup>1</sup>, Thomas Lu<sup>1</sup>, Ross K. Snider<sup>1,†</sup> & Li Liang<sup>1</sup>

It has been well documented that neurons in the auditory cortex of anaesthetized animals generally display transient responses to acoustic stimulation, and typically respond to a brief stimulus with one or fewer action potentials<sup>1–5</sup>. The number of action potentials evoked by each stimulus usually does not increase with increasing stimulus duration<sup>1,5–7</sup>. Such observations have long puzzled researchers across disciplines and raised serious questions regarding the role of the auditory cortex in encoding ongoing acoustic signals. Contrary to these long-held views, here we show that single neurons in both primary (area A1) and lateral belt areas of the auditory cortex of awake marmoset monkeys (*Callithrix jacchus*) are capable of firing in a sustained manner over a prolonged period of time, especially when they are driven by their preferred stimuli. In contrast, responses become more transient or phasic when auditory cortex neurons respond to non-preferred stimuli. These findings suggest that when the auditory cortex is stimulated by a sound, a particular population of neurons fire maximally throughout the duration of the sound. Responses of other, less optimally driven neurons fade away quickly after stimulus onset. This results in a selective representation of the sound across both neuronal population and time.



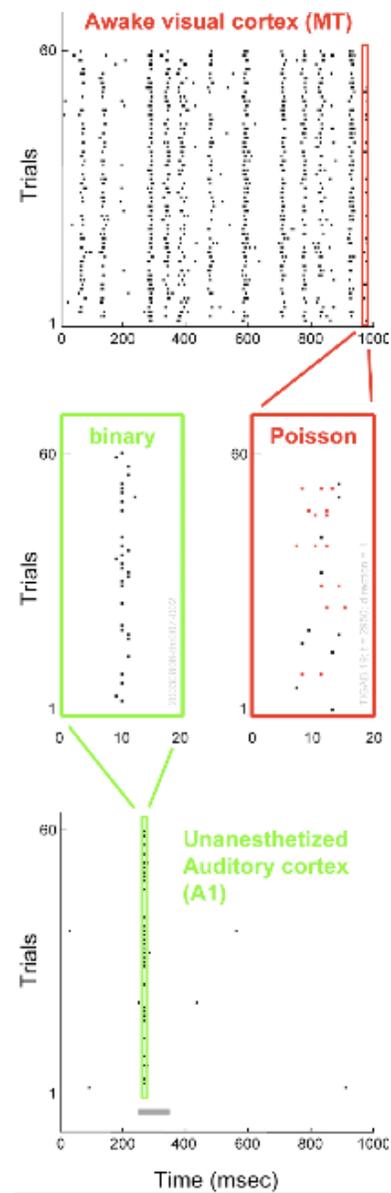
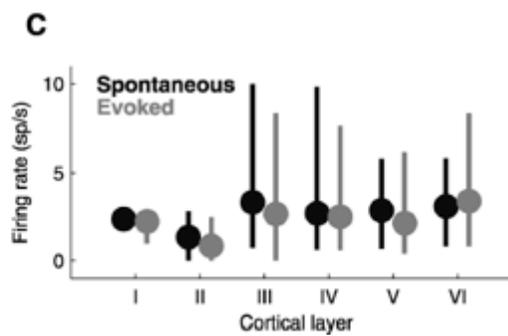
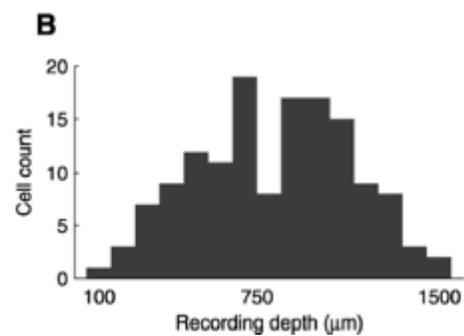
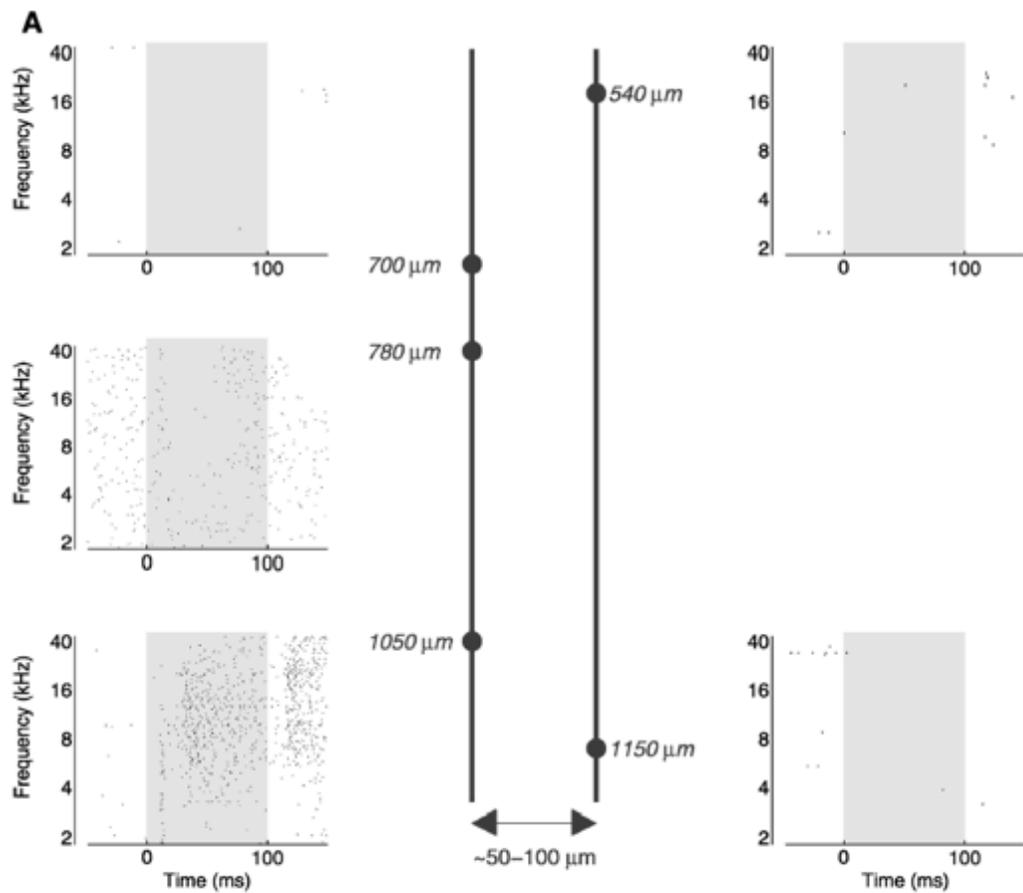
# Sparse Representation of Sounds in the Unanesthetized Auditory Cortex

Tomáš Hromádka<sup>1</sup>, Michael R. DeWeese<sup>2</sup>, Anthony M. Zador<sup>3,\*</sup>

**1** Cold Spring Harbor Laboratory, Watson School of Biological Sciences, Cold Spring Harbor, New York, United States of America, **2** Department of Physics and Helen Wills Neuroscience Institute, University of California, Berkeley, California, United States of America, **3** Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America

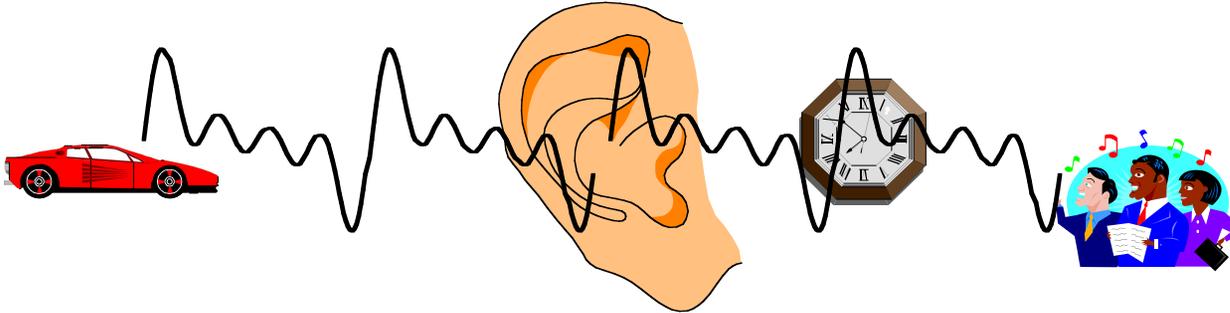
**How do neuronal populations in the auditory cortex represent acoustic stimuli? Although sound-evoked neural responses in the anesthetized auditory cortex are mainly transient, recent experiments in the unanesthetized preparation have emphasized subpopulations with other response properties. To quantify the relative contributions of these different subpopulations in the awake preparation, we have estimated the representation of sounds across the neuronal population using a representative ensemble of stimuli. We used cell-attached recording with a glass electrode, a method for which single-unit isolation does not depend on neuronal activity, to quantify the fraction of neurons engaged by acoustic stimuli (tones, frequency modulated sweeps, white-noise bursts, and natural stimuli) in the primary auditory cortex of awake head-fixed rats. We find that the population response is sparse, with stimuli typically eliciting high firing rates (>20 spikes/second) in less than 5% of neurons at any instant. Some neurons had very low spontaneous firing rates (<0.01 spikes/second). At the other extreme, some neurons had driven rates in excess of 50 spikes/second. Interestingly, the overall population response was well described by a lognormal distribution, rather than the exponential distribution that is often reported. Our results represent, to our knowledge, the first quantitative evidence for sparse representations of sounds in the unanesthetized auditory cortex. Our results are compatible with a model in which most neurons are silent much of the time, and in which representations are composed of small dynamic subsets of highly active neurons.**

Citation: Hromádka T, DeWeese MR, Zador AM (2008) Sparse representation of sounds in the unanesthetized auditory cortex. PLoS Biol 6(1): e16. doi:10.1371/journal.pbio.0060016



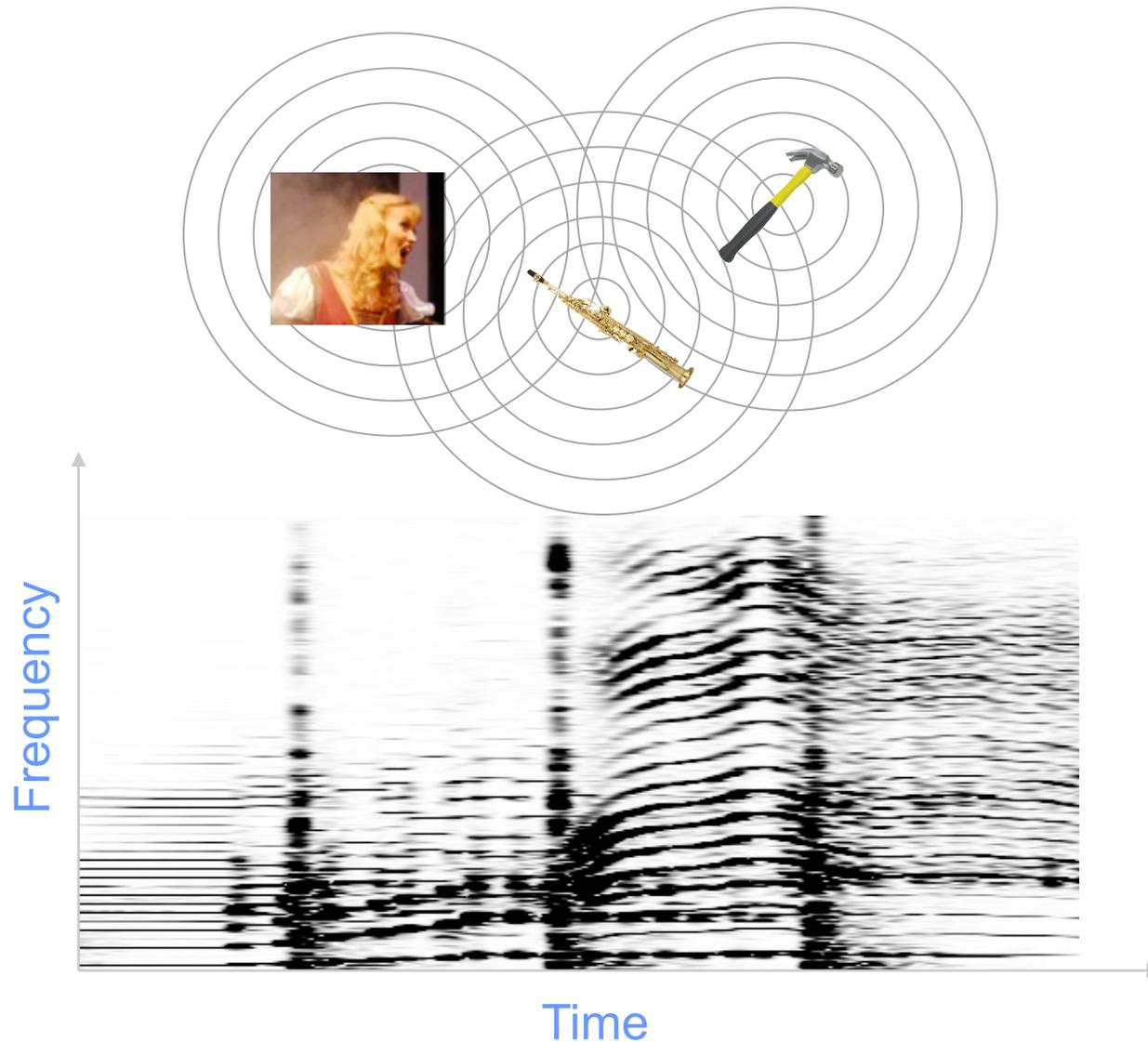
# Auditory Scene Analysis

*(sound source determination)*

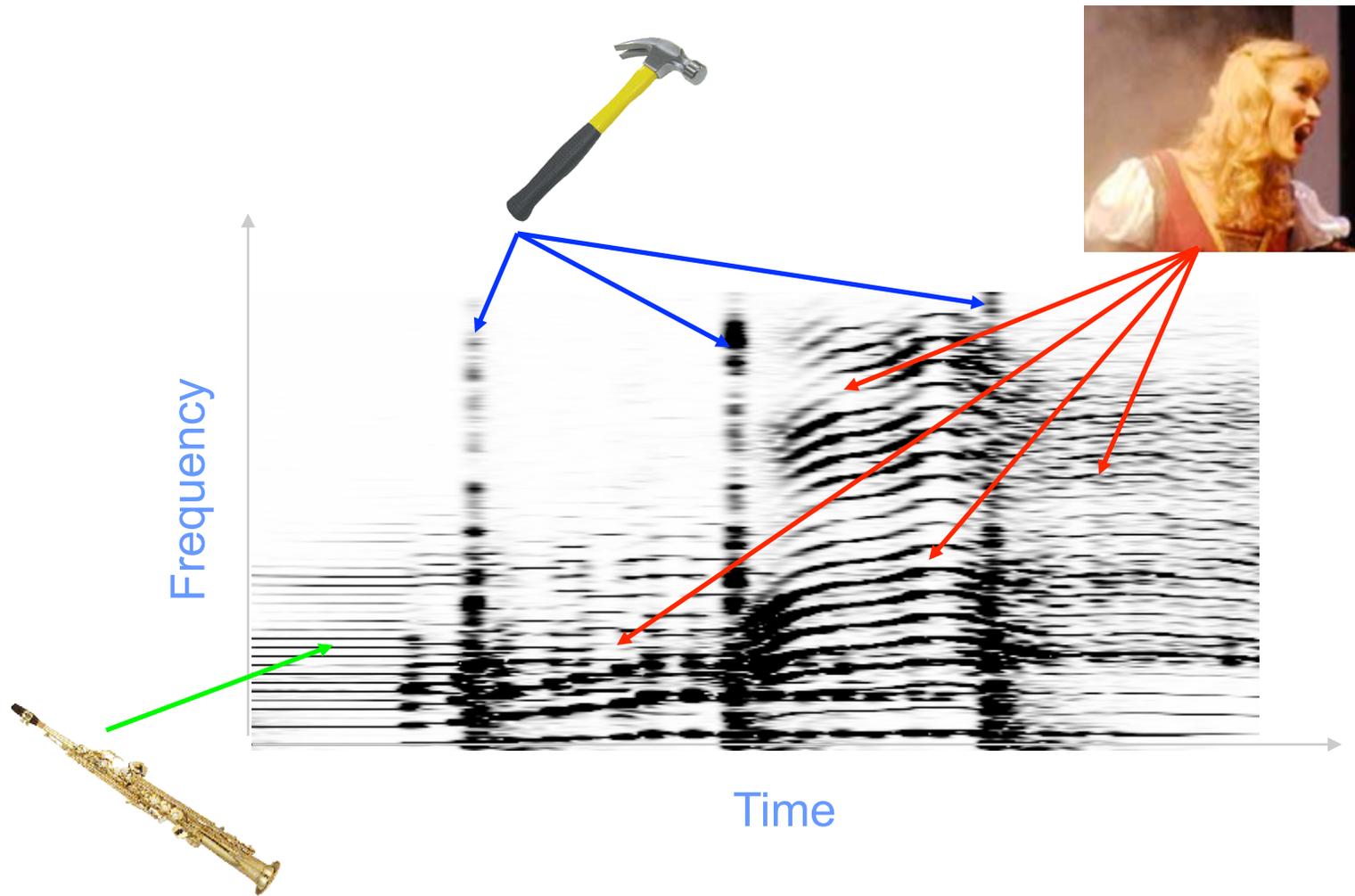


- Sounds from multiple sources arrive at the auditory periphery as one complex sound field. Individual sources are then resolved from spectro-temporal patterns.
- A single source can stimulate a wide region of the cochlea, with multiple sources simultaneously stimulating the same peripheral frequency channels.
- Determining the source of a sound depends on analytic processing to segregate spectral components according to source.

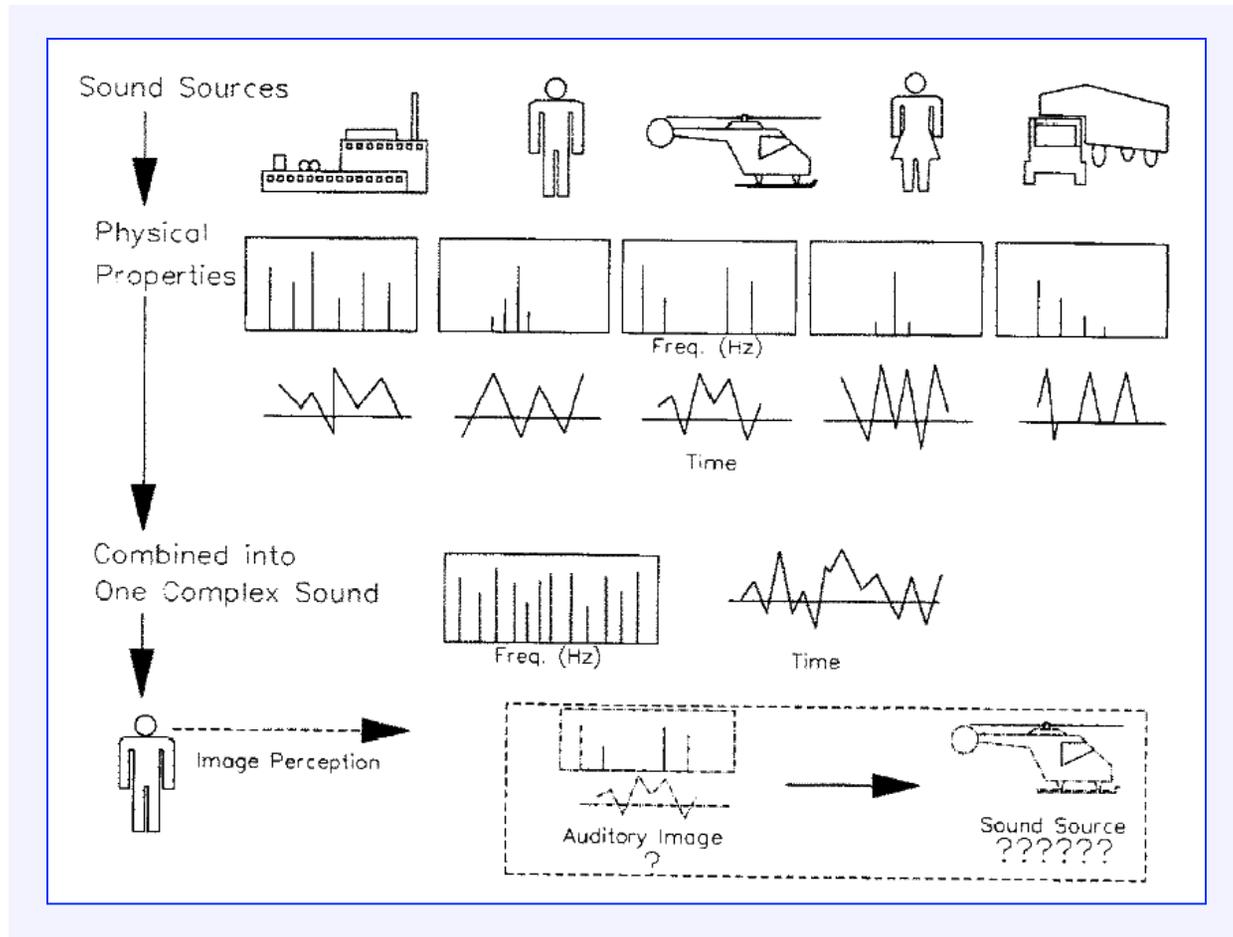
# Waves from environmental sound sources mingle before reaching the ears



# Auditory system disentangles the sources



# Yost calls this “Sound Source Determination”



# Grouping Principles

*Auditory scene analysis is well described by many of the Gestalt principles recognized in visual scene analysis (Bregman)*

## Proximity:

- Tones close in frequency will group together, constraining the extent of frequency jumps and the number of streams.
- Tones with similar timbre will group together.
- Speech sounds of similar pitch will tend to be heard from the same speaker.
- Sounds from different locations are harder to group together across time than those from the same location.

## Common fate:

- Sounds from a common source tend to start and stop at the same time and change in amplitude or frequency together (vibrato).
- A single component is easy to hear out if it is the only one to change in a complex.

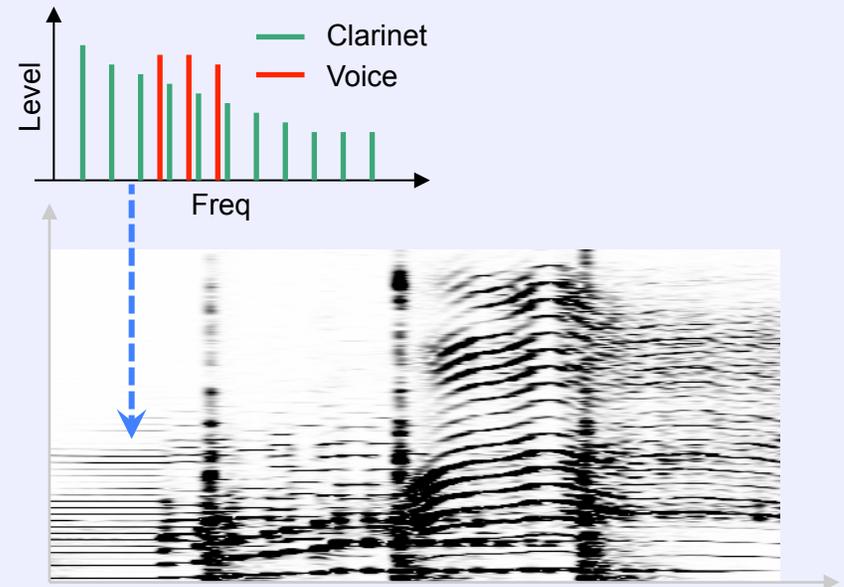
## Good continuation:

- Abrupt discontinuities in frequency or pitch, can give the impression of a different sound source.

# Perceptual Grouping

## Simultaneous

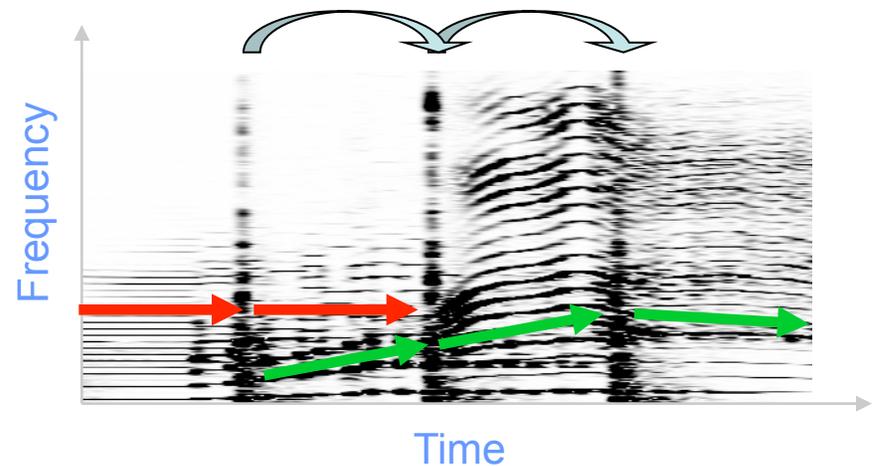
- Separating frequency components that arrive simultaneously from multiple sources (e.g., the segregation of a single voice).
- Simultaneous grouping features include spectral separation, harmonicity, onset/offset asynchrony, co-modulation, etc.



## Sequential

### 'Streaming'

- Deciding which group of components at one time is a continuation of a previous group. (e.g., connecting the parts of the same melody).
- Temporal order judgment is good within one stream but bad between streams.

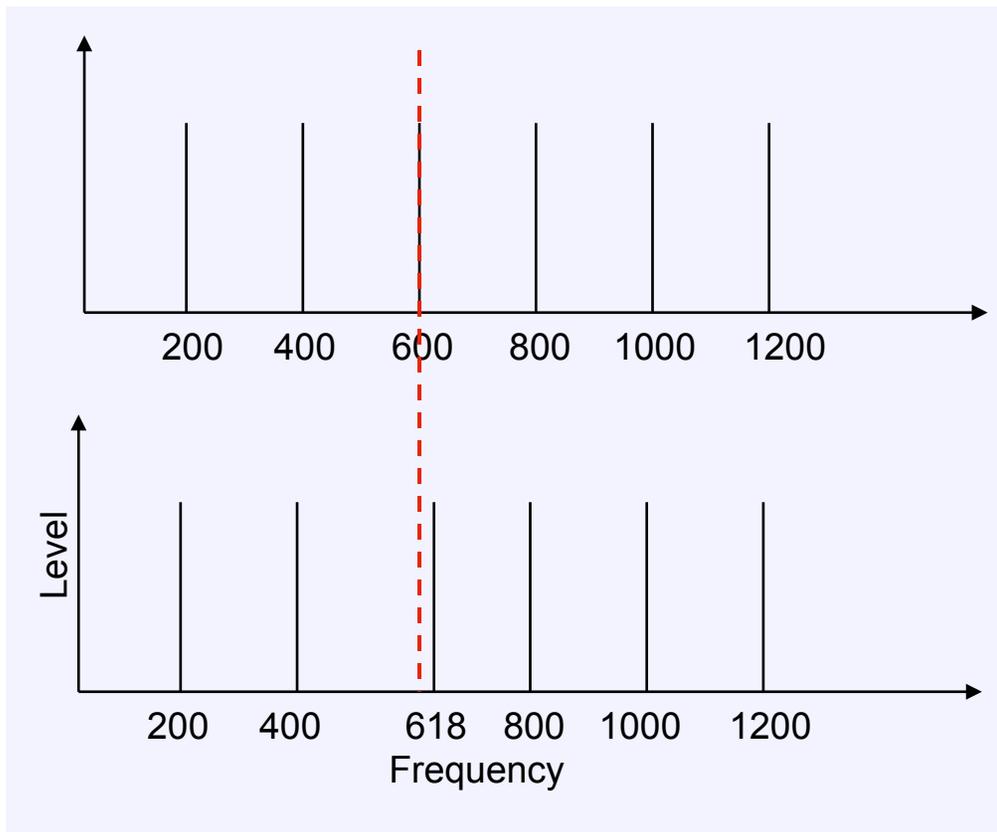


# Harmonic mistuning promotes segregation

*If a harmonic is mistuned by > 2-3%, it stands out perceptually*

*(Moore et al., 1985, 1986; Hartmann et al., 1990)*

**stimulus**



**percept**

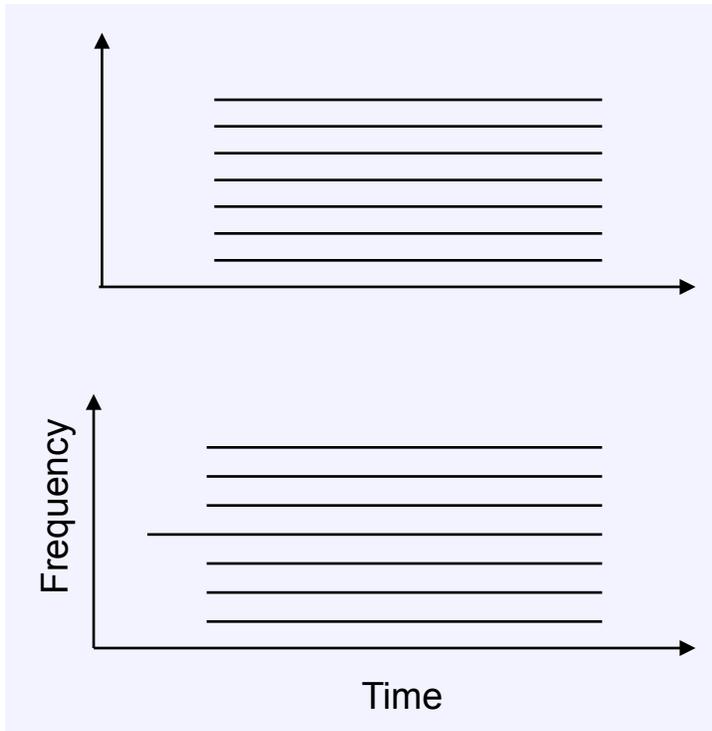
1 sound  
*pitch: 200 Hz*

2 sounds  
*Harmonic (pitch: 200 Hz)*  
*Pure tone (618 Hz)*

# Onset asynchrony promotes perceptual segregation

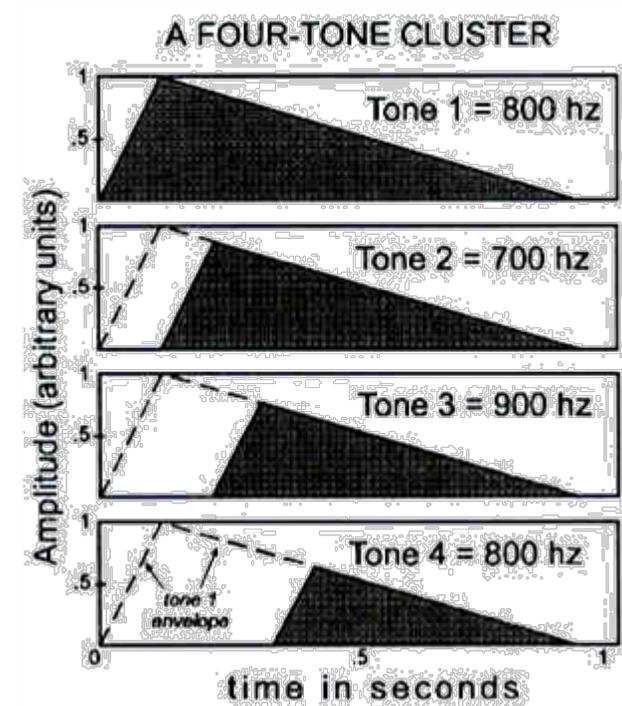
stimulus

percept



*1 sound*

*2 sounds*

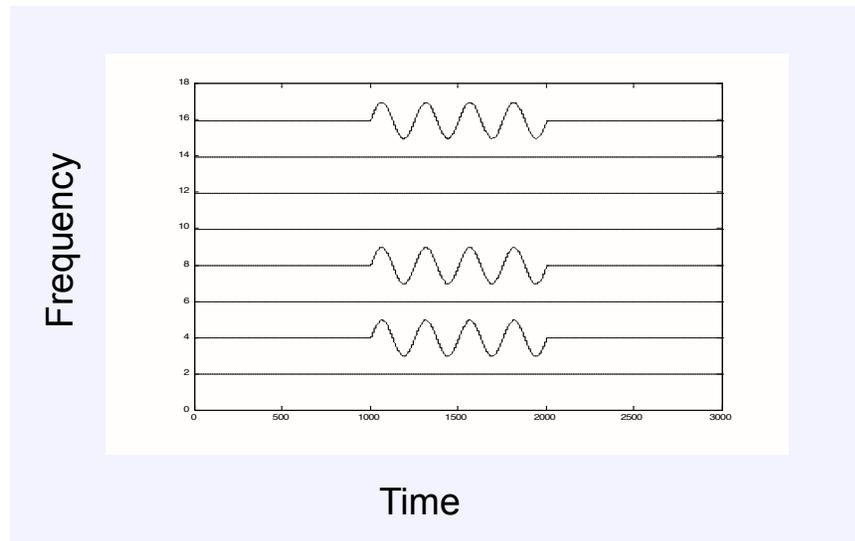


# Co-modulation promotes perceptual segregation

*Coherent FM promotes the fusion of harmonics*

Darwin et al. (1994)

stimulus



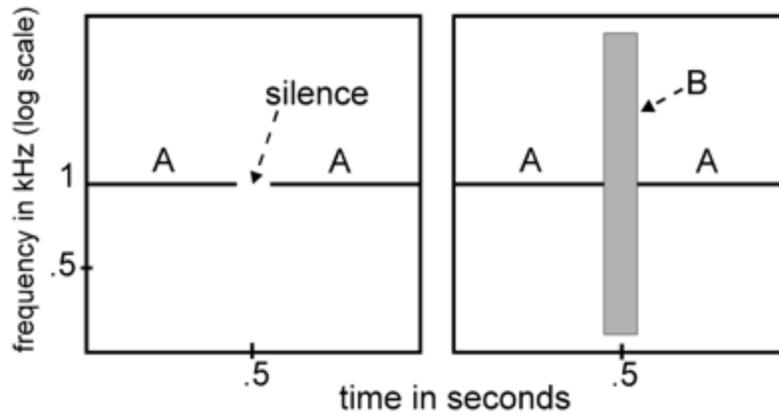
percept

*2 sounds*



# Perceptual continuity

*Sound interrupted by a noise that masks it, can appear continuous* — the auditory system interpolates across the mask, creating the illusion of continuity.



perceptual continuation of a gliding tone through a noise burst

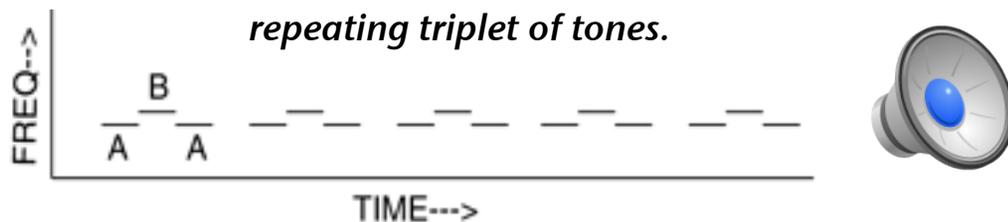


picket fence effect with speech

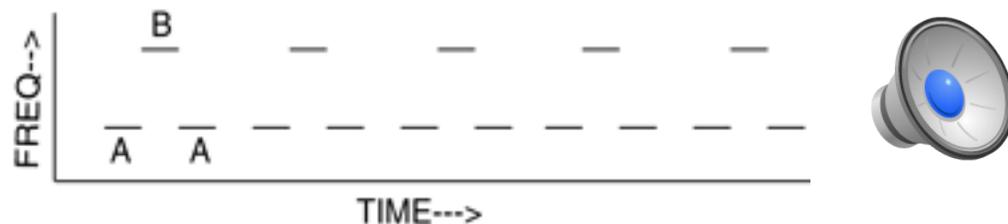


# Auditory streaming

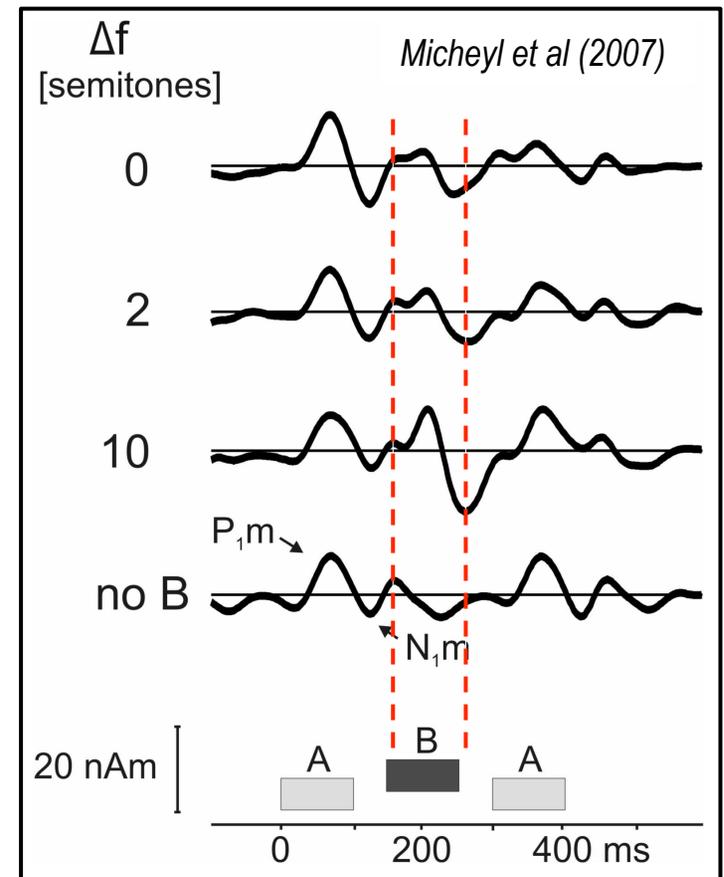
If the tone frequencies are similar, the percept is a single stream of sound with a galloping rhythm.



But if the "A" and "B" tones are far apart the gallop disappears, and two streams are heard.



12 semitones = 1 octave

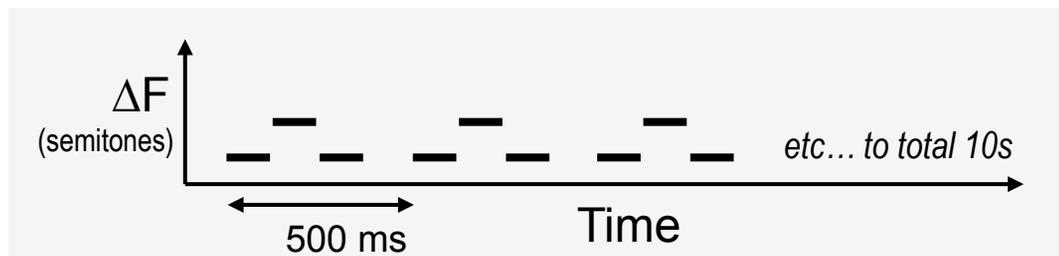


Average source waveforms from auditory cortex of 14 human listeners in response to ABA tone triplets with different A-B frequency separations

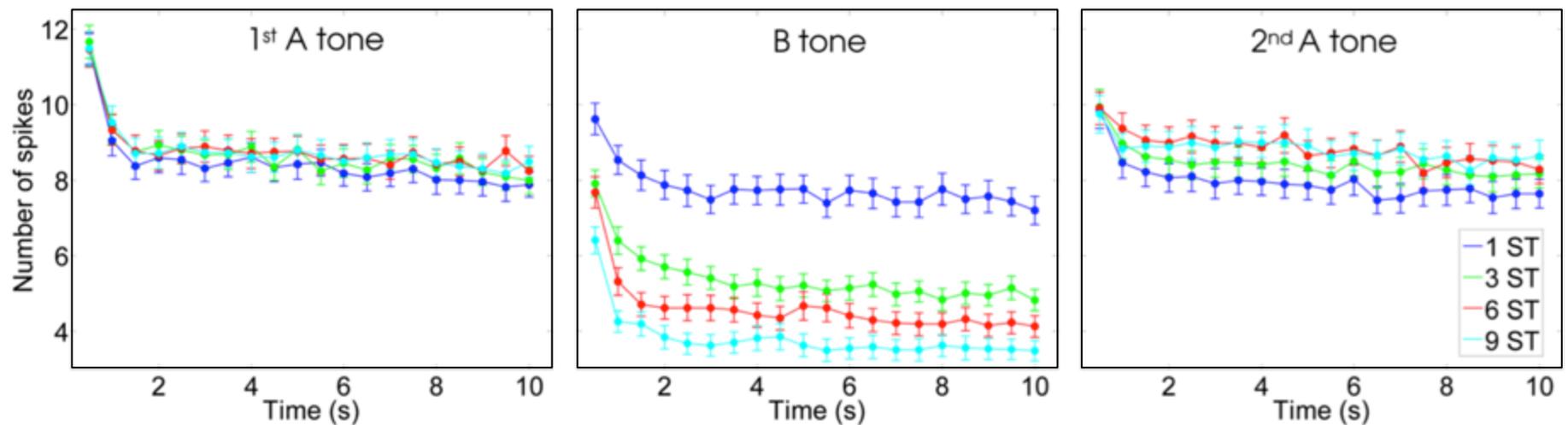
# Auditory streaming

*Responses in awake rhesus macaque auditory cortex*

Number of spikes evoked in A1 during each of 20 successive ABA-triplet sequences.



*Segregation builds up over time, and depends on the AB frequency separation*



*Adapted from Micheyl et al (2007)*

# Auditory streaming

## *Caveats on interpretation of neural data*

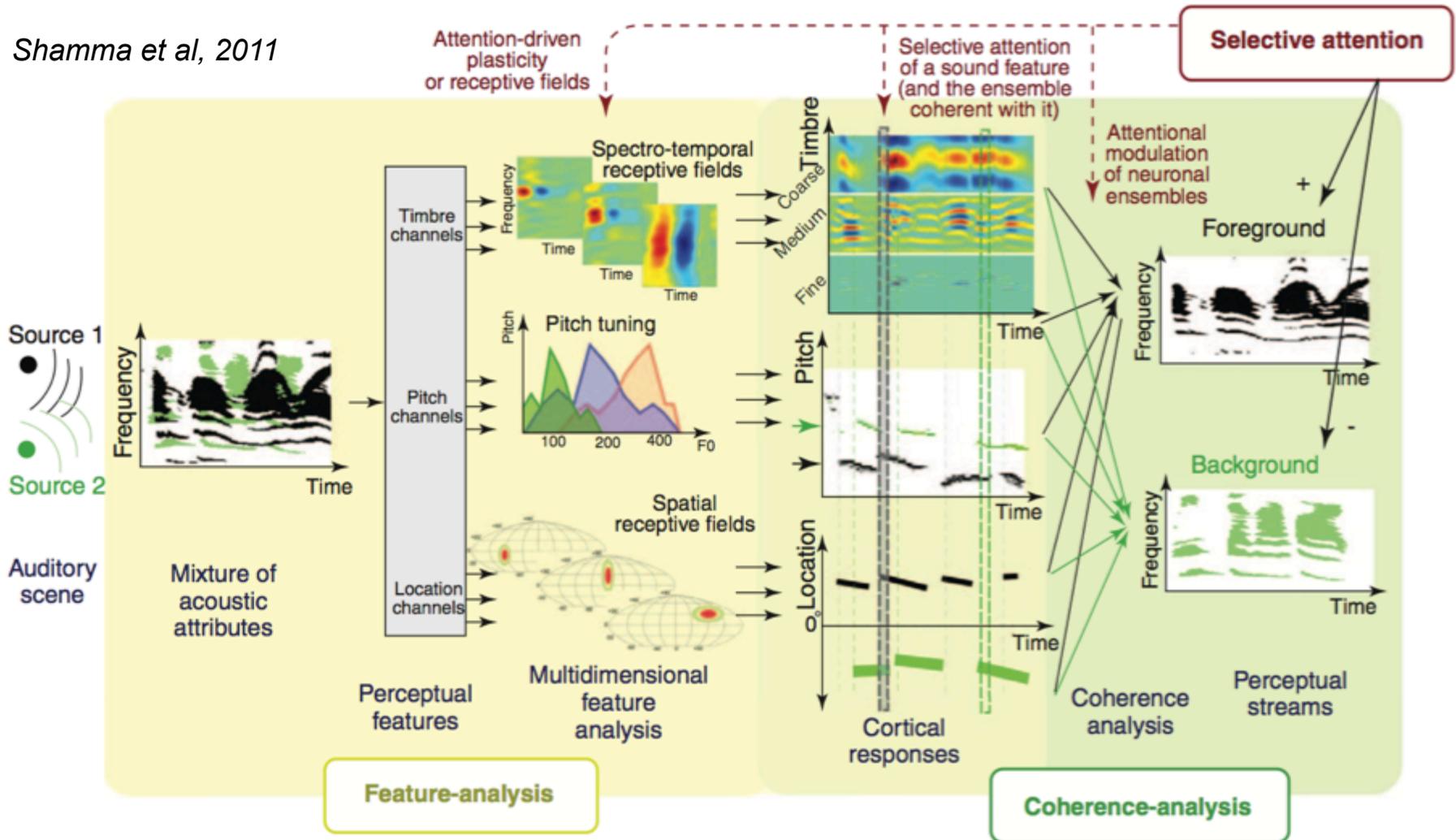
- Certain neural properties in auditory cortex, including tonotopicity and multi-second adaptation are consistent with known perceptual characteristics of streaming, but that doesn't mean that these cortical properties *determine* the streaming percepts. *e.g., those properties have also been observed below cortex.*
- Much of the current evidence is a comparison of neural responses obtained in one species with psychophysical data collected in another.
- Most neural studies have used only pure tones, but perceptual segregation occurs with many stimulus types, including sounds that activate overlapping tonotopic channels (e.g. Grimault et al., 2001). The sensitivity of A1 neurons to temporal cues may be important in stream segregation.
- Psychoacoustic experiments have shown that attention can influence the build-up of stream segregation (Carlyon et al 2001, Cusack et al 2004), but neural studies have yet to address this issue.

# Neural Mechanisms for Stream Segregation?

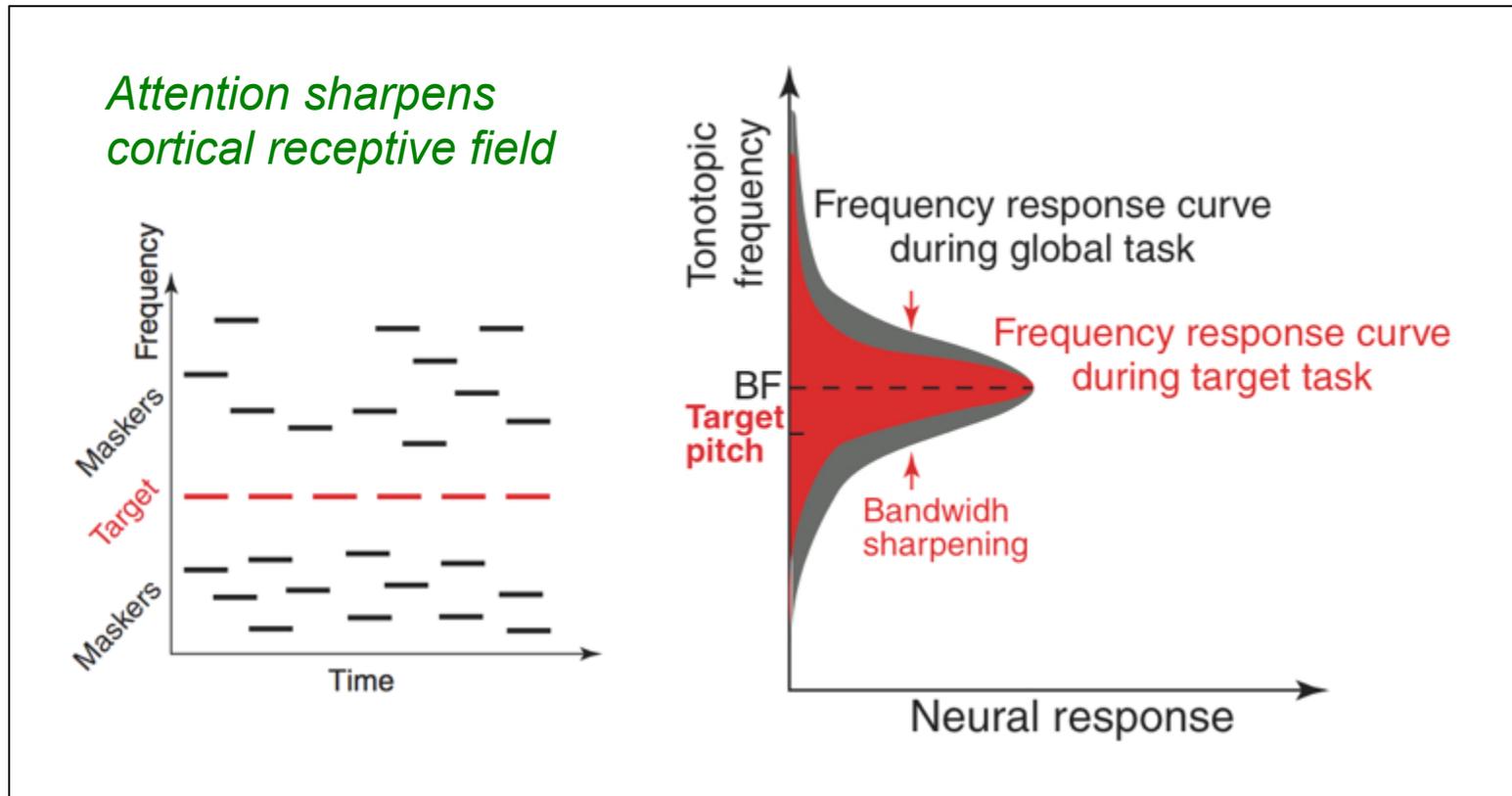
- **TEMPORAL COHERENCE** of neural responses to sounds perceptually unifies a stream, distinct from other responses.

- **ATTENTION** initiates binding, modulating neural maps and/or temporal coherence patterns of relevant acoustic features.

*Shamma et al, 2011*



# Attention in Auditory Streaming



- Frequency tuning is sharper in a ferret's auditory cortex when the ferret attends to a repeating target tone than when he listens to the entire sound mixture.
- The sharper RF enhances segregation of the target from the background of the maskers.

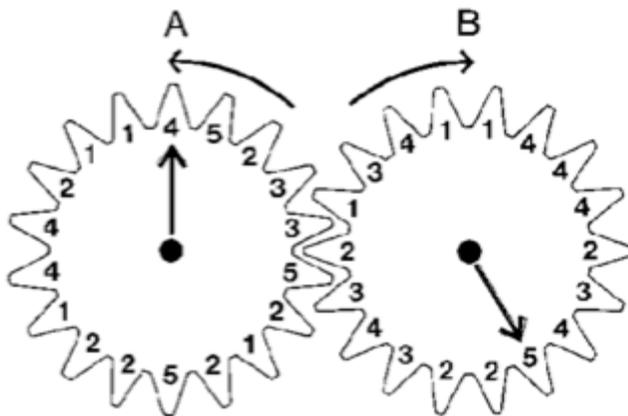
# Auditory streaming

D I S M T E R L A O C D T Y O R S

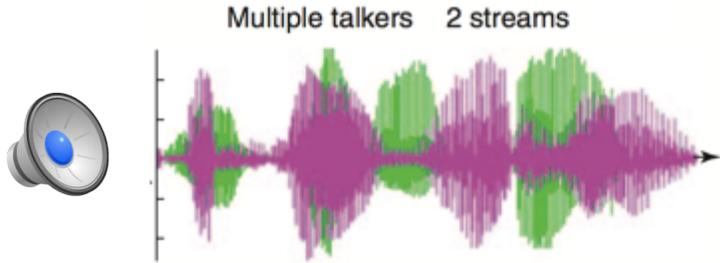
D I S <sup>M</sup>T <sup>E</sup>R <sup>L</sup>A <sup>O</sup>C <sup>D</sup>T <sup>Y</sup>O R S



*Segregation of a melody from interfering tones. On each repetition the pitch of the melody is raised (by two semitones). At some point the listener will be able to segregate the melody from the distractor tones.*



*Streaming in Ugandan xylophone music: two musicians play note sequences equally spaced in time (isochronous). Interleaving of these cycles creates high and low perceptual streams with irregular rhythms.*



# Cues for Sound Segregation

## **Spectral separation**

- Sometimes two or more sounds having different spectral components can be resolved by frequency selectivity.

## **Harmonicity**

- A sound consisting of harmonics is usually perceived as a single pitch equal to the fundamental frequency even if the fundamental is absent in the sound spectrum.
- Fusion of spectral components into an image or entity is described as a pitch.

## **Profile analysis & Timbre**

- Sound spectra are characterized by their intensity variation as a function of frequency.
- The auditory system is sensitive to changes in spectral profile despite variation in overall sound level.
- Timbre allows discrimination of one sound from another when both have the same pitch, loudness, and duration. Different spectra give rise to different timbres.

## **Spectral Modulation**

- Changing a frequency modulation or modulating a complex set of harmonics facilitates segregation.

## **Temporal separation & Asynchrony**

- Multiple sound sources are often asynchronous (onset or offset) & temporally complex, aiding segregation.

## **Spatial separation**

- IID and ITD are cues for segregation, particularly with competing signals. *e.g., cocktail party effect*
- Increasing separation between masker and signal improves discrimination. *e.g., multiple talkers*

## **Temporal modulation**

- Coherence of change among spectral components may help the auditory system to identify those components as originating from a single source (*can be AM or FM*).
- A coherently modulated cue band offers relief from masking (*comodulation masking release*).
- A modulated masker can impair detection of a modulated signal (*modulation detection interference*).

## **Informational masking**

- The inverse relationship between pattern complexity and target resolution is called informational masking.
- Variation in target location in a pattern, total number of elements in the pattern, and stimulus uncertainty affect the information content of a sequence.