Encoding The Pitch Of Sounds Using Synchrony Receptive Fields

Jonathan Laudanski* and Romain Brette.

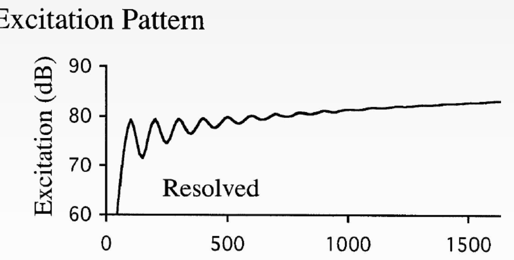
Equipe Audition, Département d'Etudes Cognitives, Ecole Normale Supérieure, 75005 Paris, France. (*email:jonathan.laudanski@ens.fr)

Introduction: 1

Pitch constitutes a major dimension of auditory perception along which periodic (or near-periodic) sounds can be organized. Although the role of pitch is essential in the perception of music or when dealing with source segregation in auditory scene, the neural mechanisms underlying pitch perception remain unclear [1]. Generally two major types of pitch models are distinguished: spectral vs temporal models. Both categories account for percepts such as the missing fundamental, the pitch shift of mistuned harmonics, etc.. but each type suffers from specific issues:

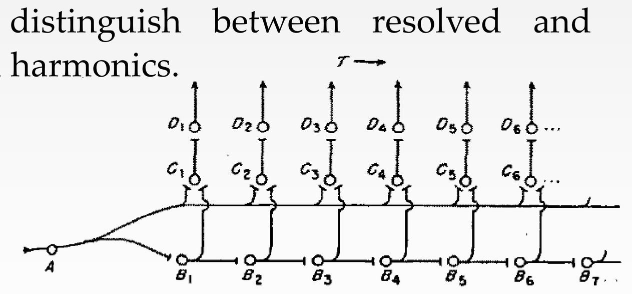
Spectral / Templates models:

- Cannot account for the pitch of unresolved harmonics.



Temporal / auto-correlation models:

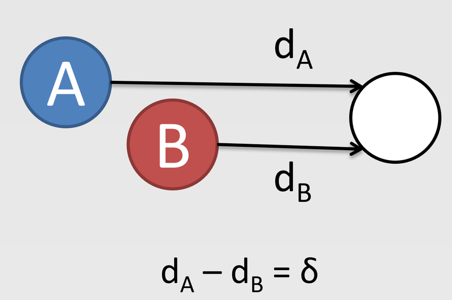
- Require large delays (30 ms) for low pitches.
- Do not distinguish between resolved and unresolved harmonics.



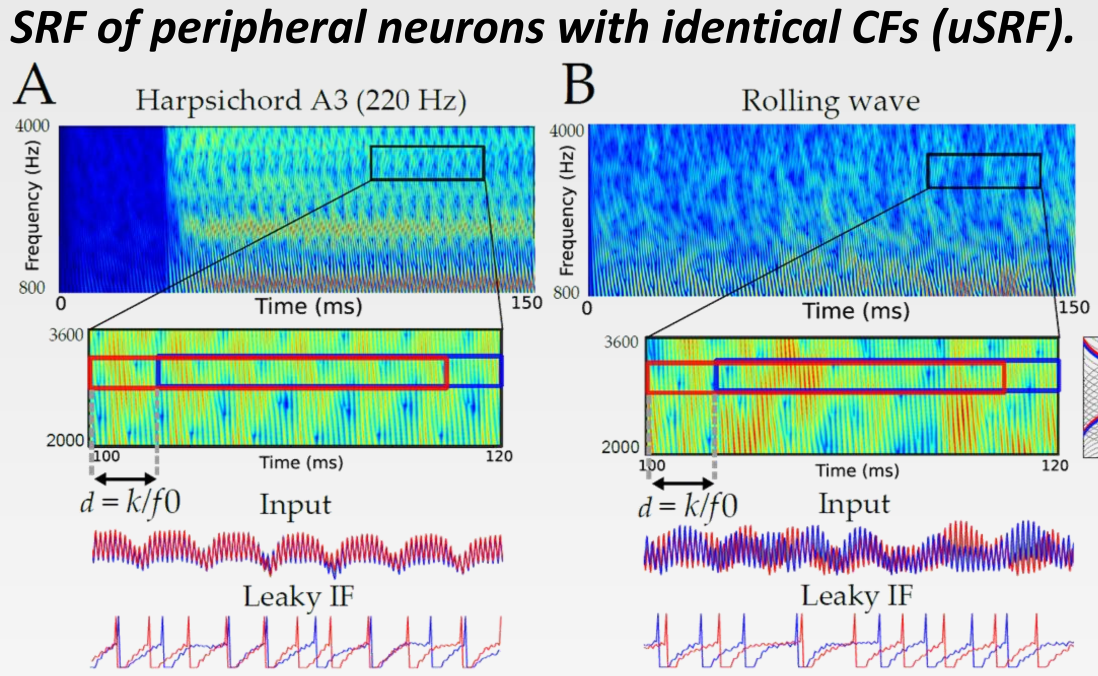
We propose that pitch is encoded in the patterns of synchrony of cochlear nucleus neurons. Resolved and unresolved harmonics produce different types of synchrony patterns. The model does not require long delays.

Synchrony Receptive Fields (SRF). 2

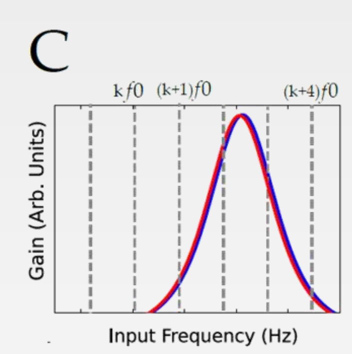
A stimulus *s(t)* belongs to the SRF of a pair (A,B) of peripheral neurons, if it produces synchronous output spikes. In this case, we have $s \otimes_t F_A = s \otimes_{t-\delta} F_B$ where *F* is the neuron's receptive field (auditory filter). More formally, the SRF of a neuron group can be defined as [2]: $SRF(F_A,F_B): \{s; s \in Ker(F_B - F_A)\}$ In the case of peripheral auditory neurons receiving input from a gammatone filter centered at CF (that is: $F_A = \Gamma_{cf}$), any periodic stimulus can only belong to two types of SRFs.



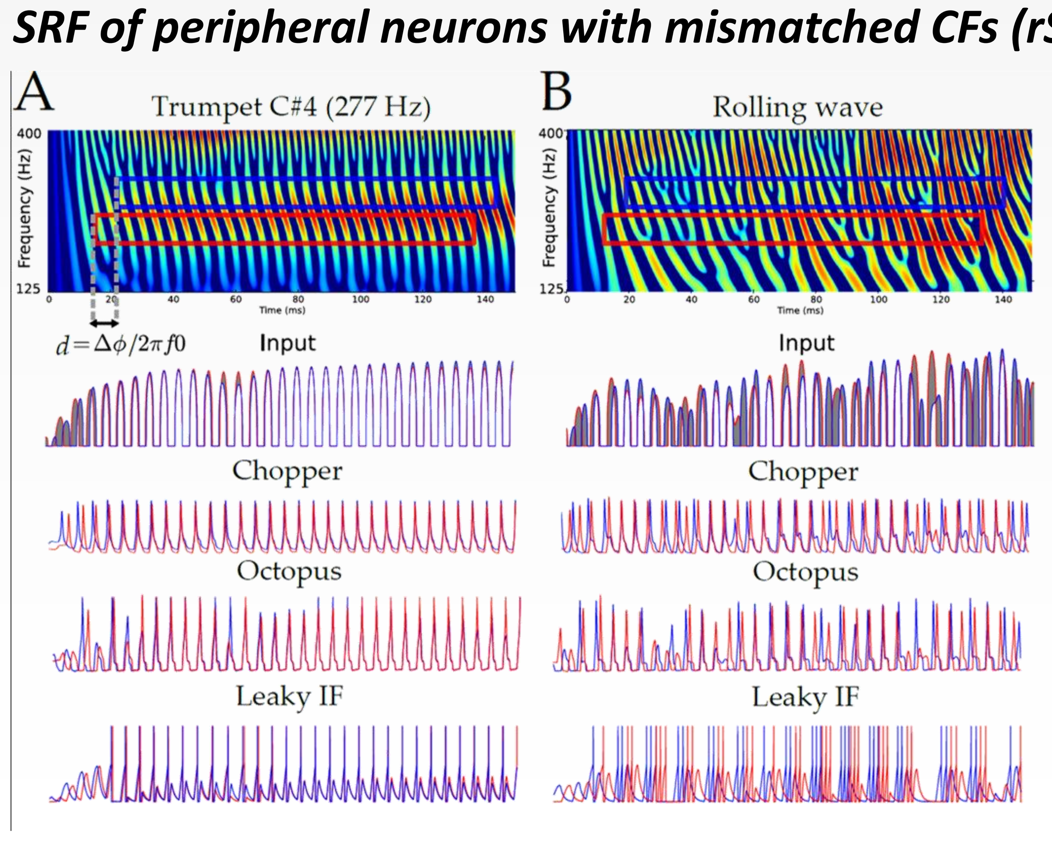
SRF of peripheral neurons with identical CFs (uSRF).

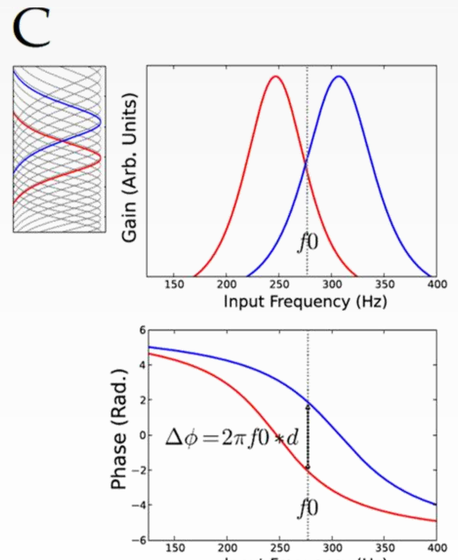


A) Synchronous responses to a set of unresolved harmonics. B) Responses to a noise are not synchronous. C) The SRF includes unresolved harmonics.



SRF of peripheral neurons with mismatched CFs (rSRF).

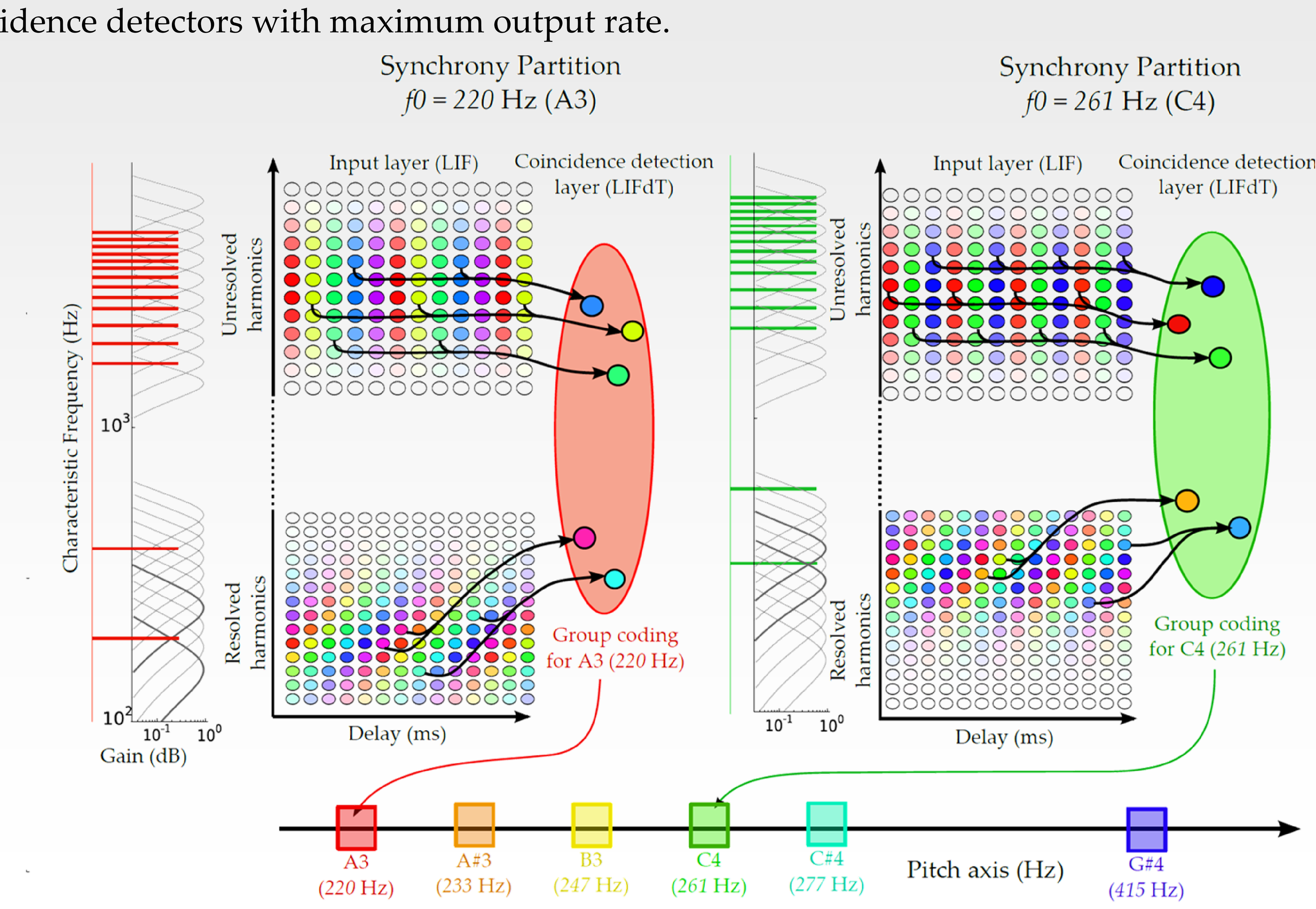




A) Synchronous responses to a trumpet note, independent of the neuron model used. B) Responses to a noise are never synchronous since inputs received by the two channels differ. C) The SRF is limited to pure tones or resolved harmonics.

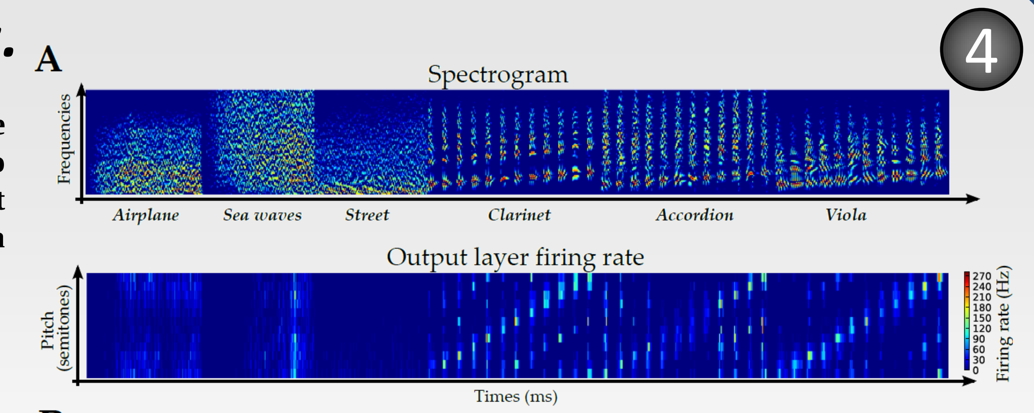
Encoding pitch using assemblies. 3

When stimulated by a periodic sound, peripheral neurons are partitioned into many synchronous groups (color-coded here by its response phase). Input neurons encoding resolved harmonics may differ in CFs (bottom groups), while those encoding unresolved harmonics will differ only by their delays (upper groups). Each partition is dependent on the stimulus period (eg: left groups vs right groups). By wiring each group onto specific coincidence detectors, one can decode pitch by finding the assembly of coincidence detectors with maximum output rate.

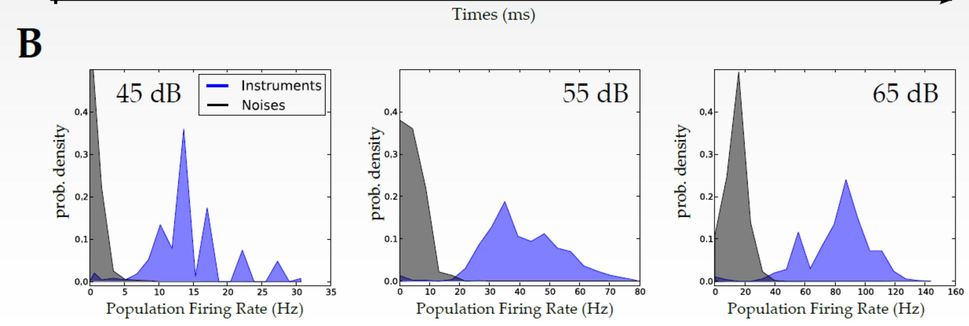


Noises vs periodic sounds. 4

A) Example of responses of the coincidence detectors (ie network output layer) to environmental noises compared to instrument notes. The network respond strongly when a pitch is present and sparsely otherwise.



B) This result generalises across sound level. Even at higher SPL input layer coincidences are more likely for periodic sounds than noises.

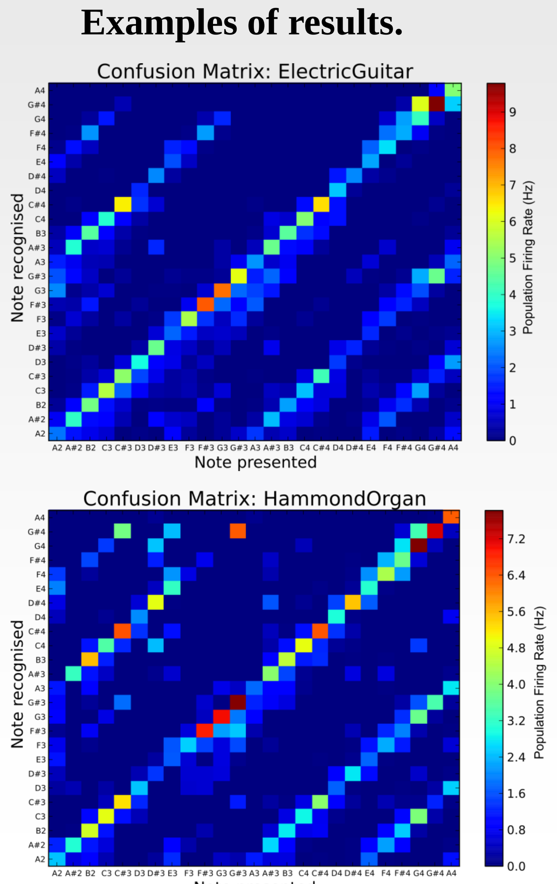


Pitch recognition task: RWC instrument database. 5

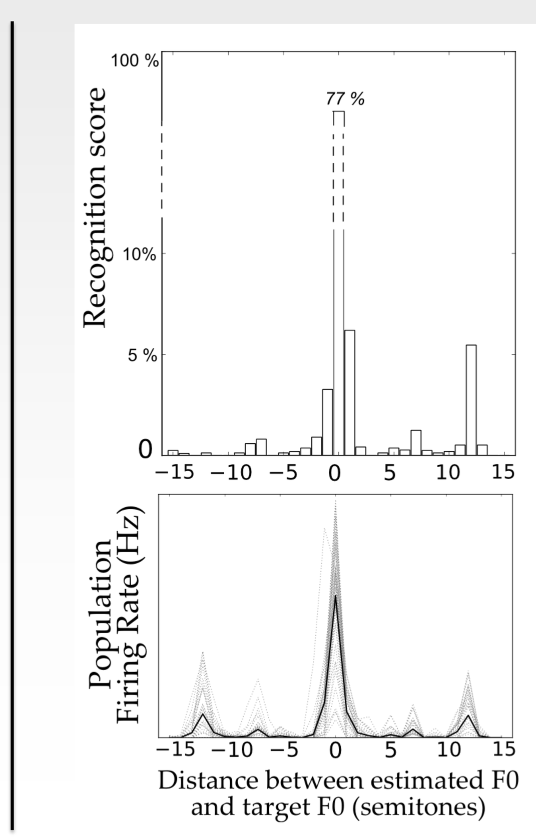
Methods:

- 500 ms sounds from the RWC instruments database [3] (762 notes from 31 instruments & voices, 5 vowels)
- Gammatone filterbank with N_{chan} = 300 channels spaced on an ERB scale with compression and half-wave rectification is used as input to noisy leaky integrate-and-fire neurons (τ_{LIF}=1.5 ms).

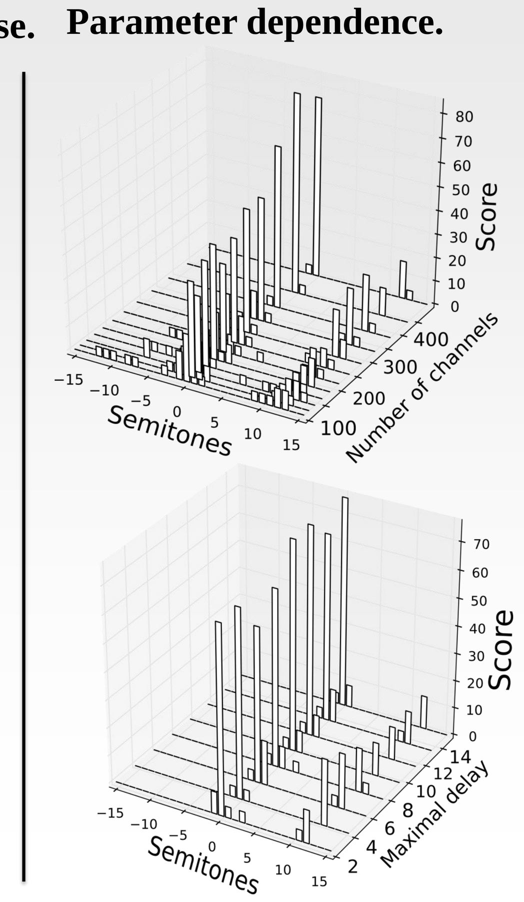
Examples of results.



Summary of results for the database.



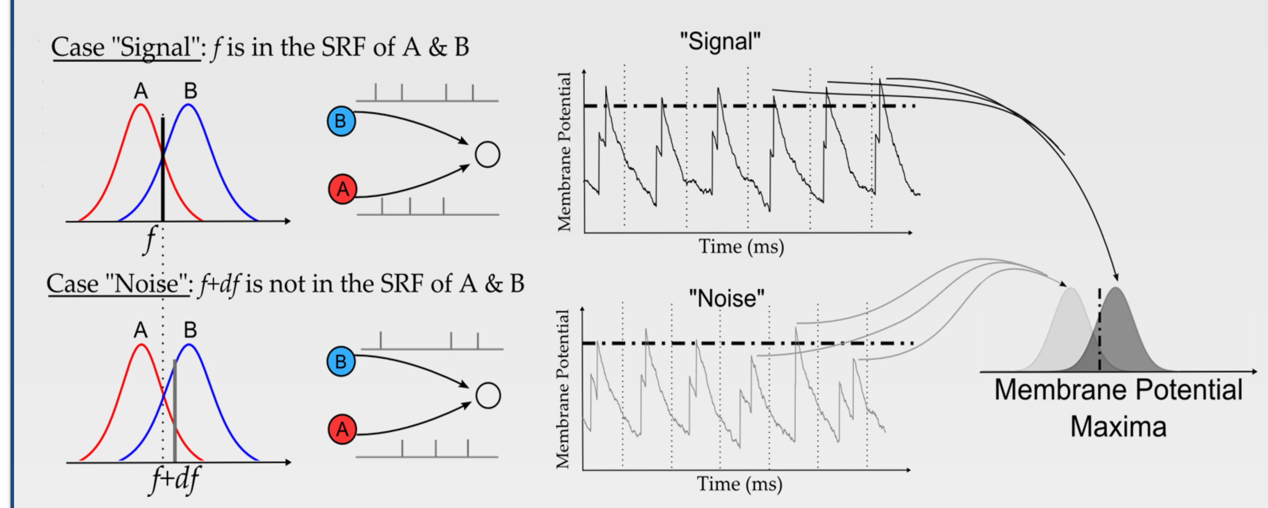
Parameter dependence.



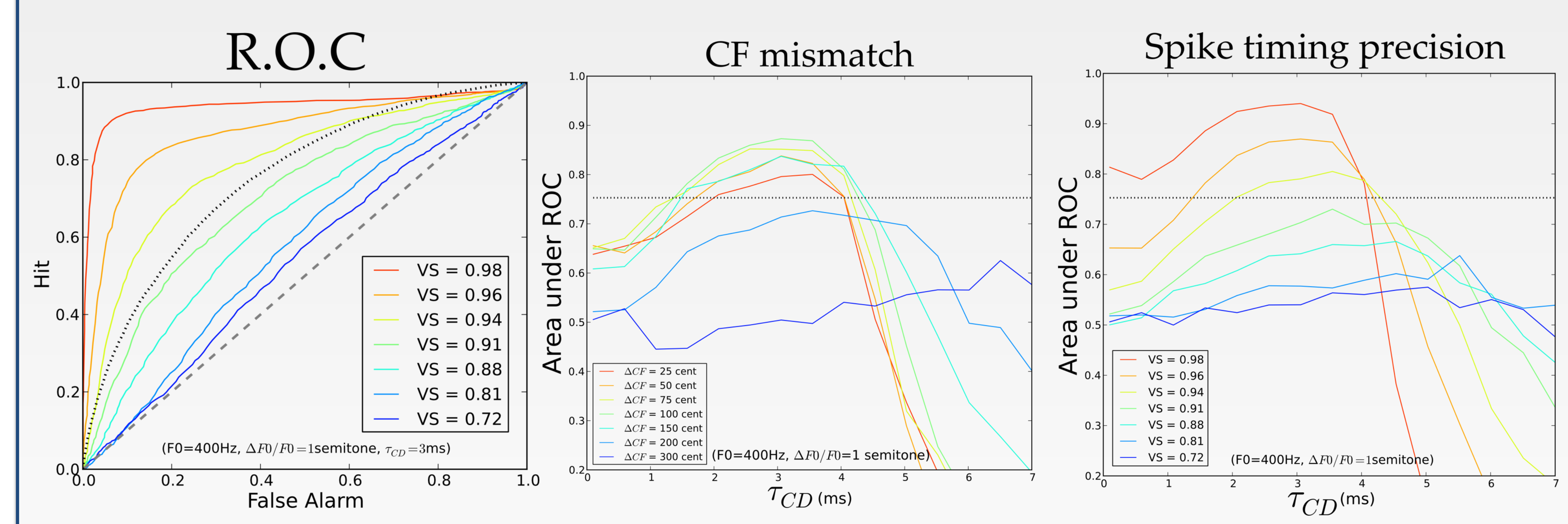
Model sensitivity: F0 discrimination by rSRF and uSRF. 6

The seminal work of Siebert [4] compared human frequency discrimination to auditory nerve data using signal detection theory (see also [5] and [6]). We use a signal detection theory to analyse the frequency discrimination ability at the level of a coincidence detector (for both uSRF and rSRF).

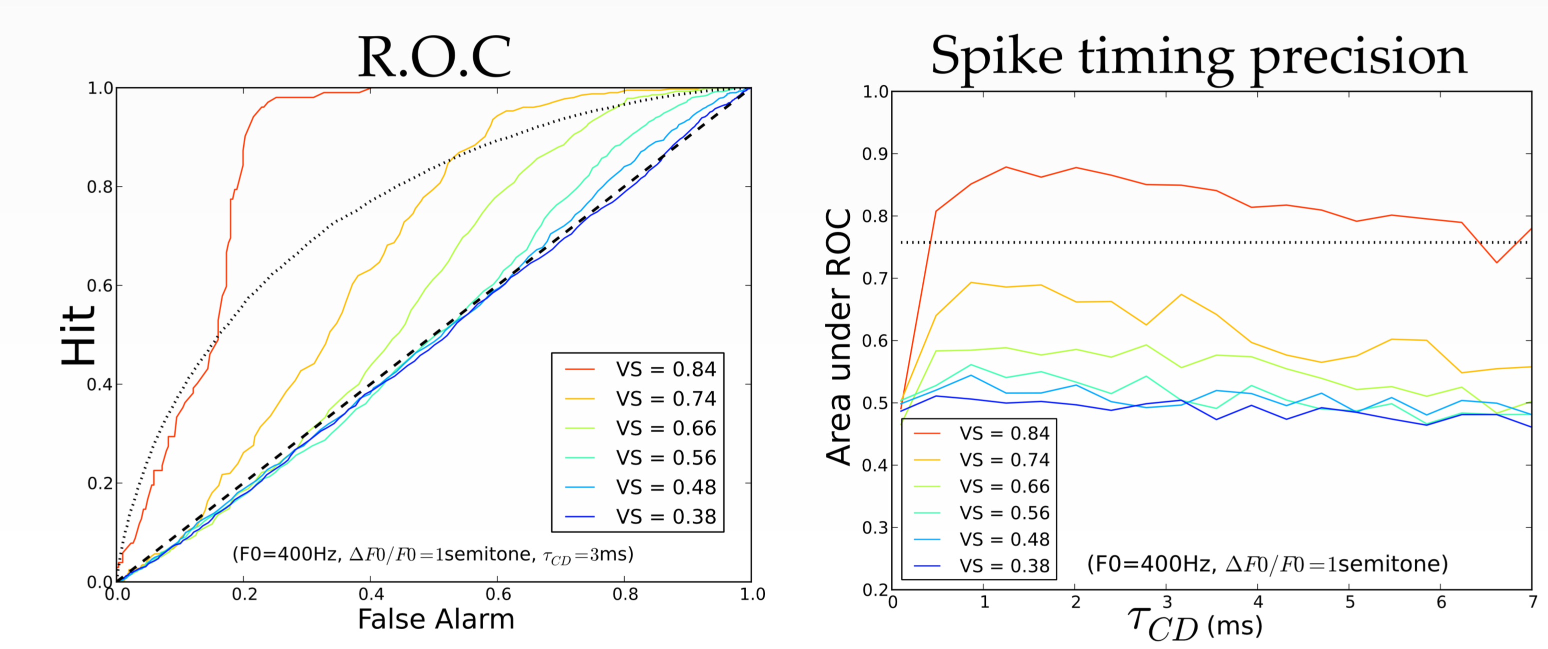
Encoding sensitivity of an rSRF:



If we consider a pair of neurons with mismatched CFs, their EPSPs arrive synchronously if the tone is in the SRF, and with some delay otherwise. Changing the threshold of their coincidence detector modifies the hit and false alarms rate. This is equivalent to moving along the R.O.C curve.



Encoding sensitivity of an uSRF:



Conclusions:

The encoding of pitch by synchronous groups of peripheral neurons proves a workable hypothesis. Decoding can be done by monitoring the firing rate of specific groups of coincidence detectors. It generalizes temporal models, with specificities:

- resolved and unresolved harmonics are encoded by different SRFs, but with the same principle;
- long delays are not necessary for resolved harmonics.

References:

[1] Winter I (2005) *The neurophysiology of Pitch. In Pitch.* Ed. Springer Handbook of Auditory Research (24), 99-146.
[2] Goodman DF and R Brette (2010). *Spike-timing-based computation in sound localization.* PLoS Comp Biol 6(11): e1000993.
[3] RWC Music Database, http://staff.aist.go.jp/m.goto/RWC-MDB/
[4] Seibert WM (1965) *Some implications of the stochastic behavior of primary auditory neurons.* Kybernetic 2, 206-215.
[5] Rieke, F.,Warland, D., de Ruyter van Stevenincк, R., & Bialek,W. (1997). *Spikes: Exploring the neural code. Cambridge, MA: MIT Press.*
[6] Heinz M, Colbrun S, and Carney L. (2001) *Evaluating Auditory Performance Limits: I. One-Parameter Discrimination Using a Computational Model for the Auditory Nerve,* Neural Computation 13, 2273–2316