Computational Perception

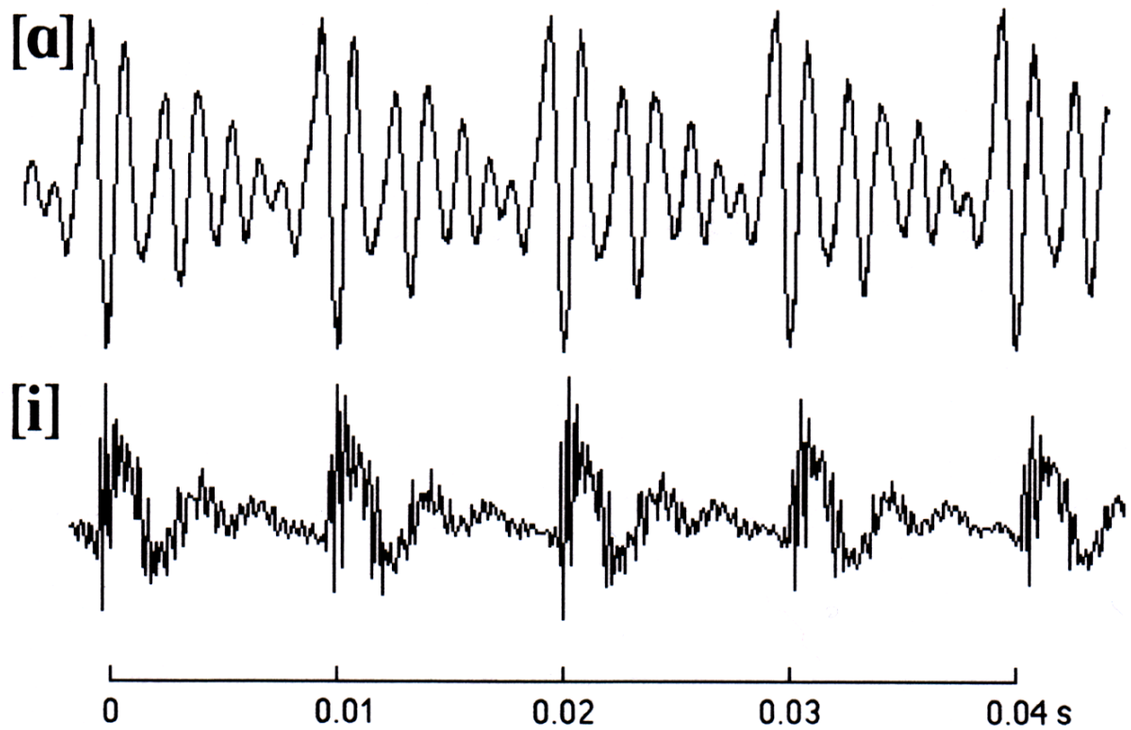
15-485/785

Auditory Structure 2

Pitch and timbre

Pitch is perceptual frequency: the frequency of a sine wave with the same perceptual pitch.

- Often equal to the rate of the repetition
- Pitch is only defined for relatively fast repetitions (>20-50 Hz)
- Slower repetitions are perceived as a distinct sequence of sounds
  - in speech: sequence of phonemes
  - in music: sequence of notes
  - within these sounds, pitch can be perceived

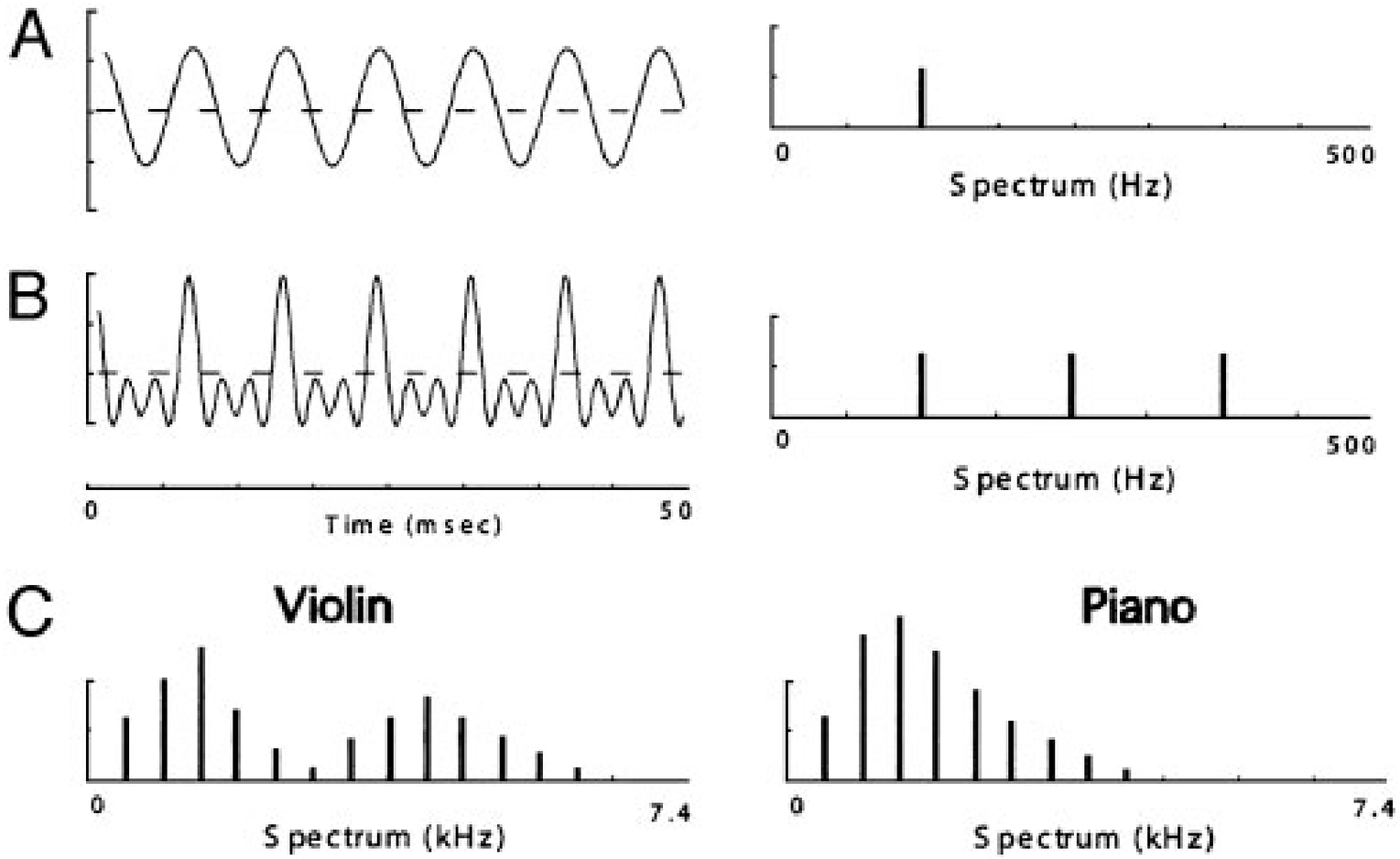


Two vowels with the same pitch

Timbre is the part that is repeated

- These characterize the resonances of the vibrating source and its acoustical enclosure.
- Loudness is not timbre, nor is position, but almost everything else is.

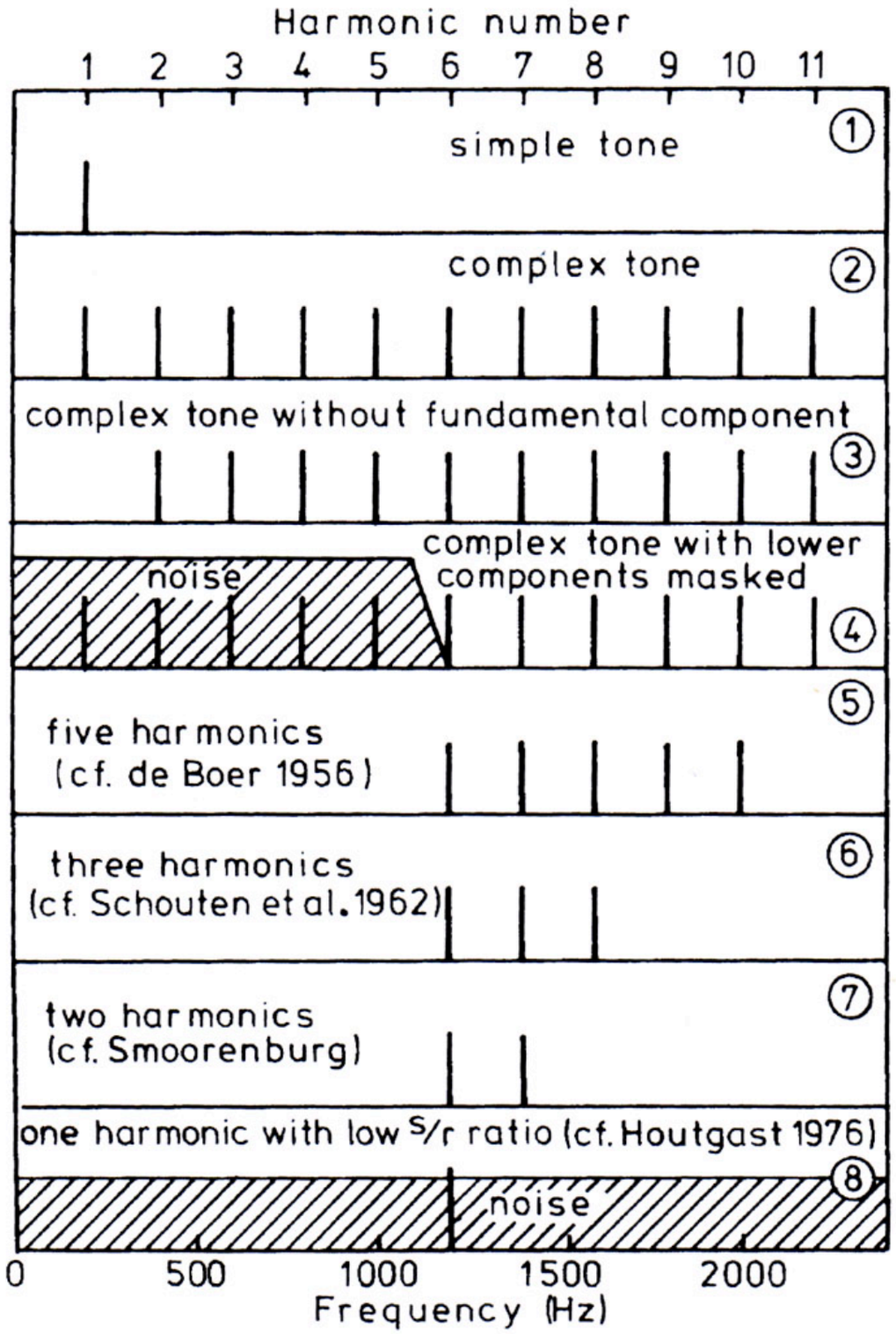
More examples



Other properties of pitch and timbre

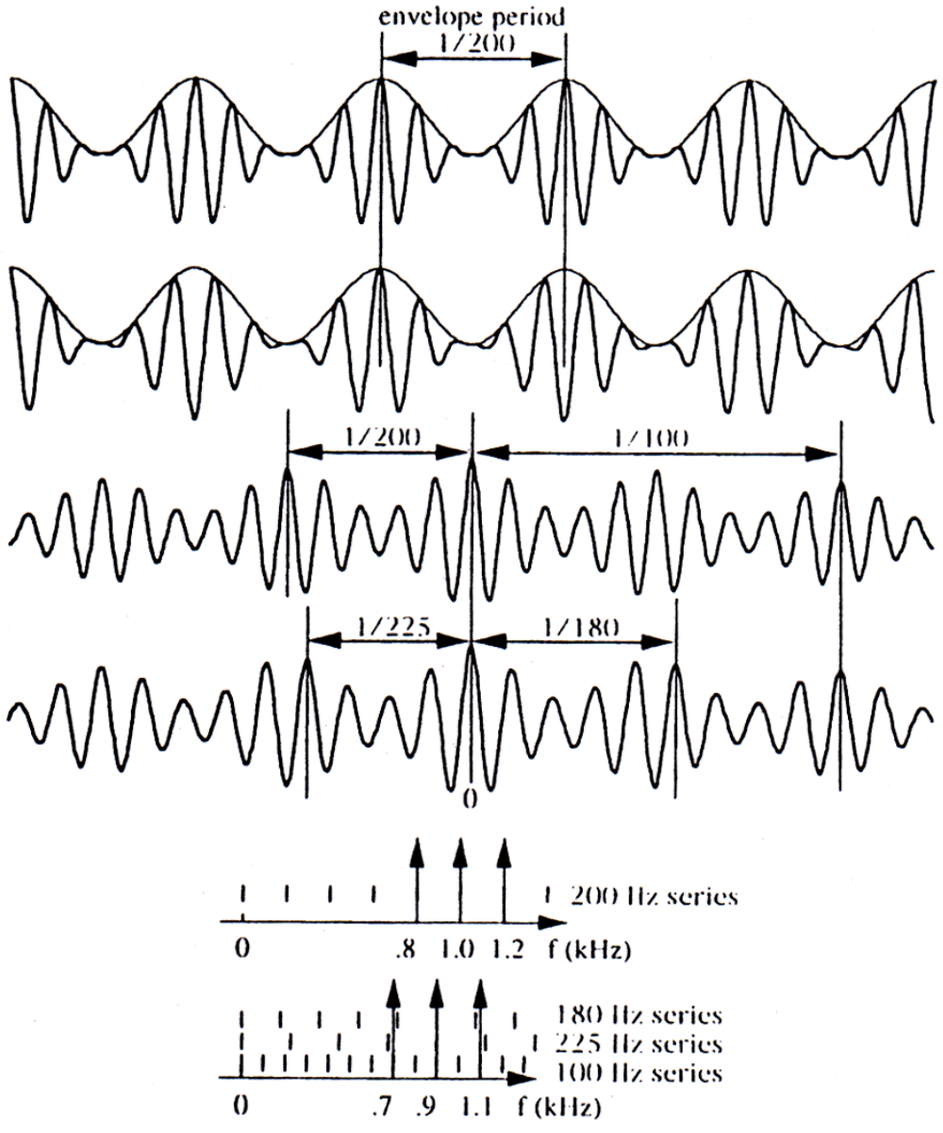
- Onsets and offsets affect timbral qualities
- Only need a limited number of repetitions
- in studies of musical note perception as few as two repetitions can elicit a clear pitch percept
- Transient sounds (ie non-repetitive) can have timbral qualities, but not pitch
- Percussive sounds (eg drums) can have fairly clear pitch without repetitive excitation because of an approximately periodic waveform
- Periodic sounds can have perceived pitch different from the repetition rate

Eight signals with the same low pitch



- clicks 200 times per second
  - has low pitch very close to 200 Hz pure tone
  - contains harmonics with frequencies of 200, 400, 600, 800, etc Hz
  - can filter to remove 200 Hz, pitch unchanged, timbre slightly different
  - can even eliminate all but 1800, 2000, 2200, low pitch remains, timbre very different
  - low pitch is present even when masked by low freq noise
- low pitch is called "residue"
  - also called "periodicity pitch", "virtual pitch", and "low pitch"
  - perception of residue pitches are what we normally hear when we listen to complex tones
- Theories of residue pitch
  - pattern recognition models
  - time interval models

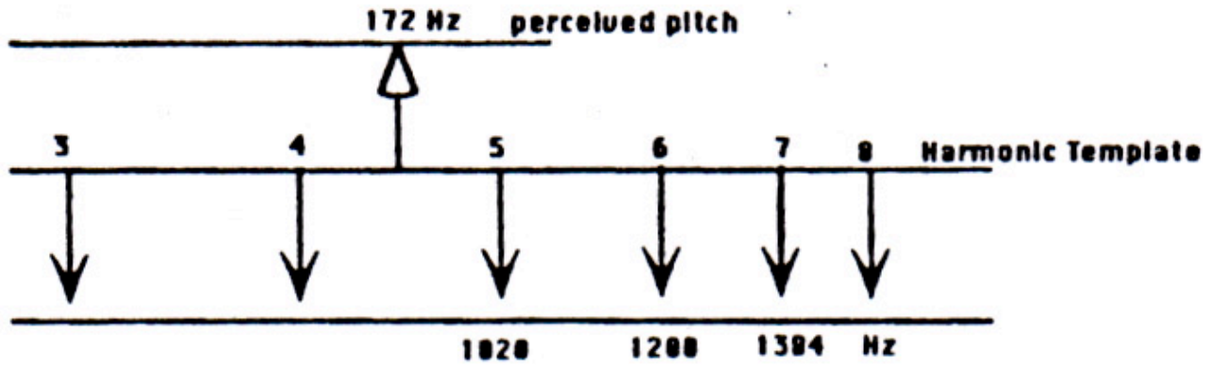
Pitch is not the frequency of the amplitude envelope



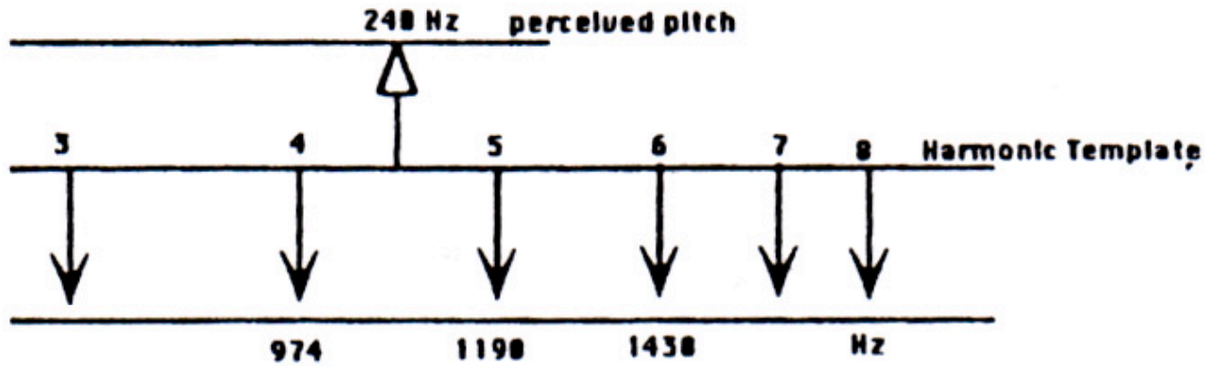
- top signal: 800/1000/1200 Hz
  - perceived pitch: 200 Hz
- second signal: 700/900/1100 Hz
  - perceived pitch: ~175 and 233 Hz
- perceived pitch is closer to fine time scale intervals

Place theories of pitch perception

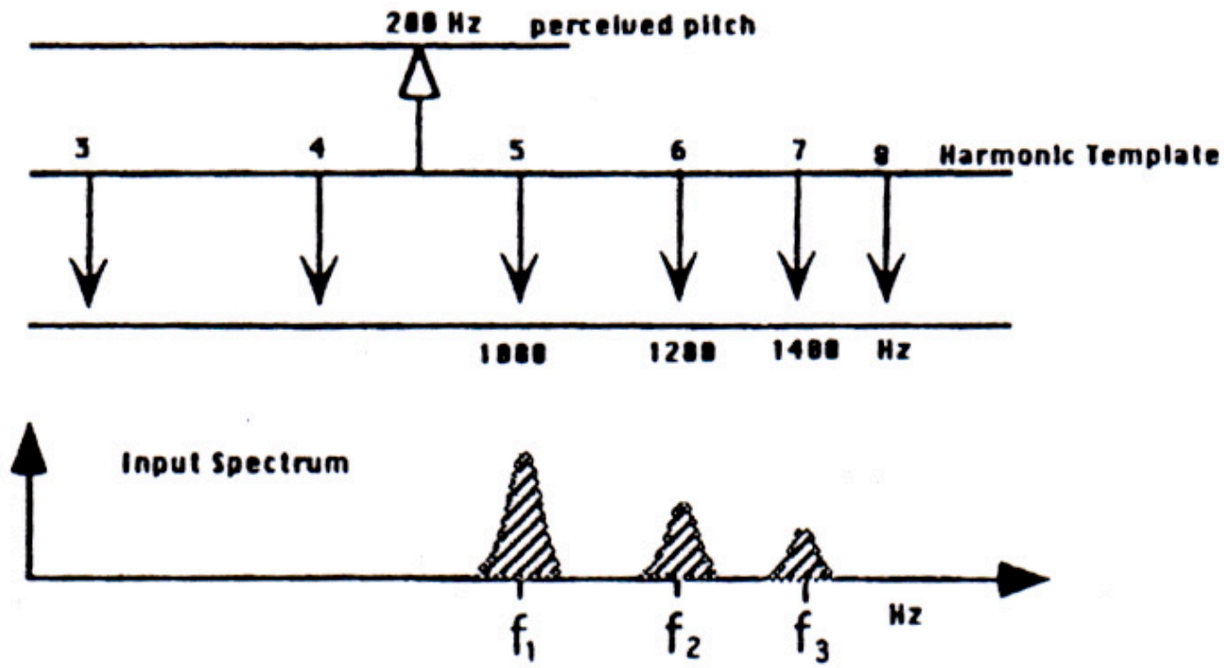
Example 3
Best match is
between the 172 Hz
harmonic template
and the input spectrum



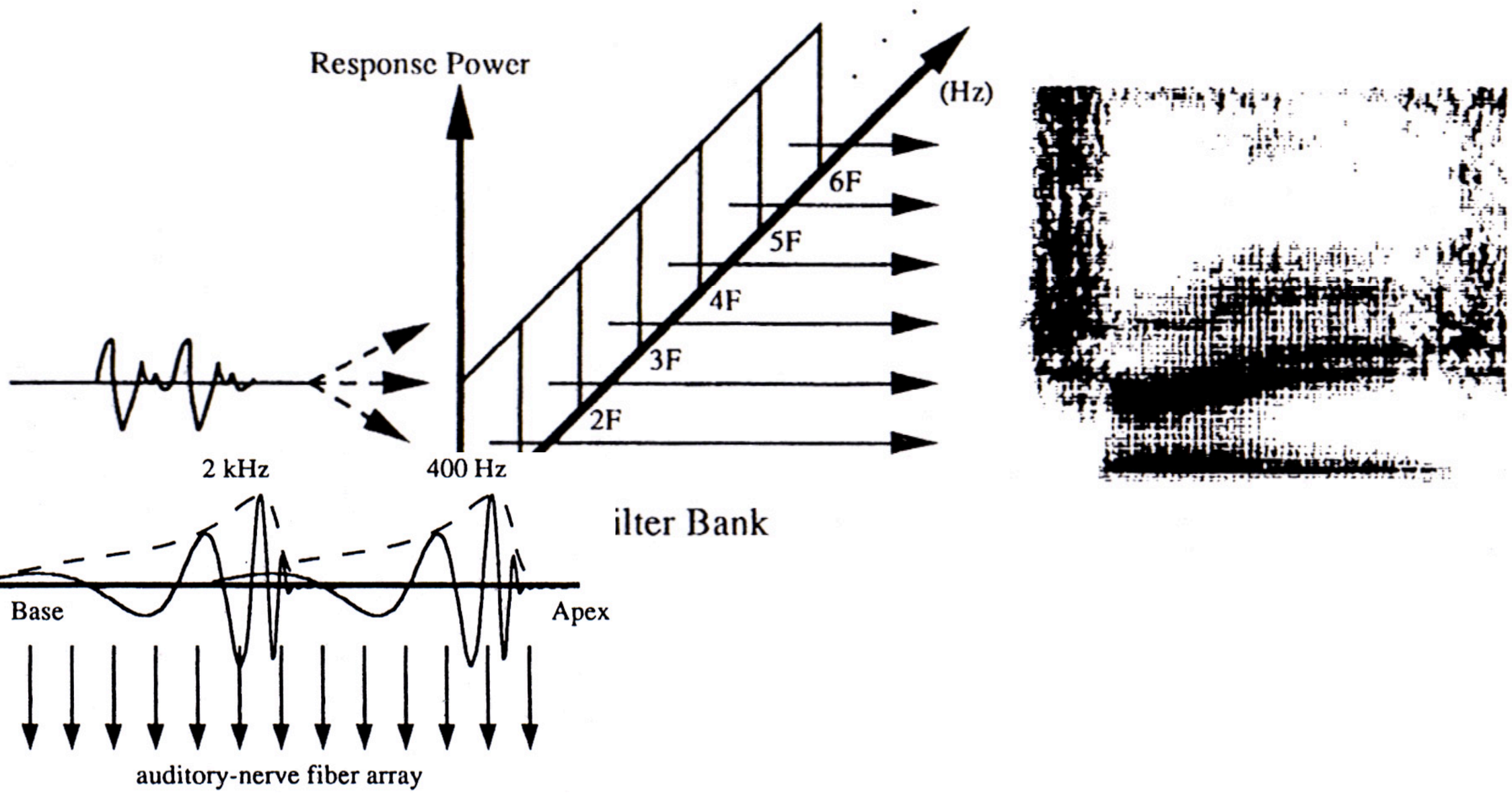
Example 2
Best match is
between the 240 Hz
harmonic template
and the input spectrum



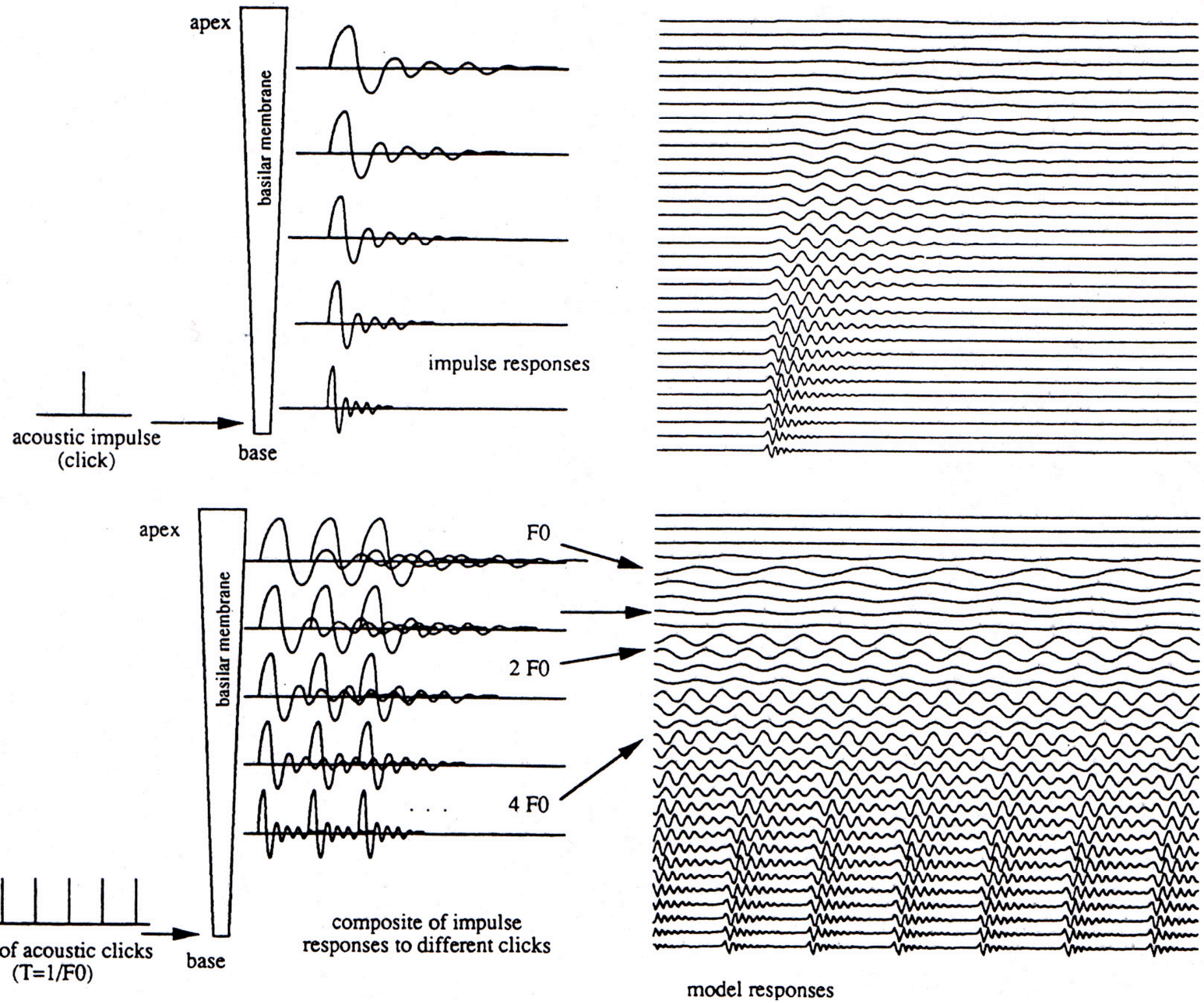
Example 1
Optimum match is
between the 200 Hz
harmonic template
and the input spectrum



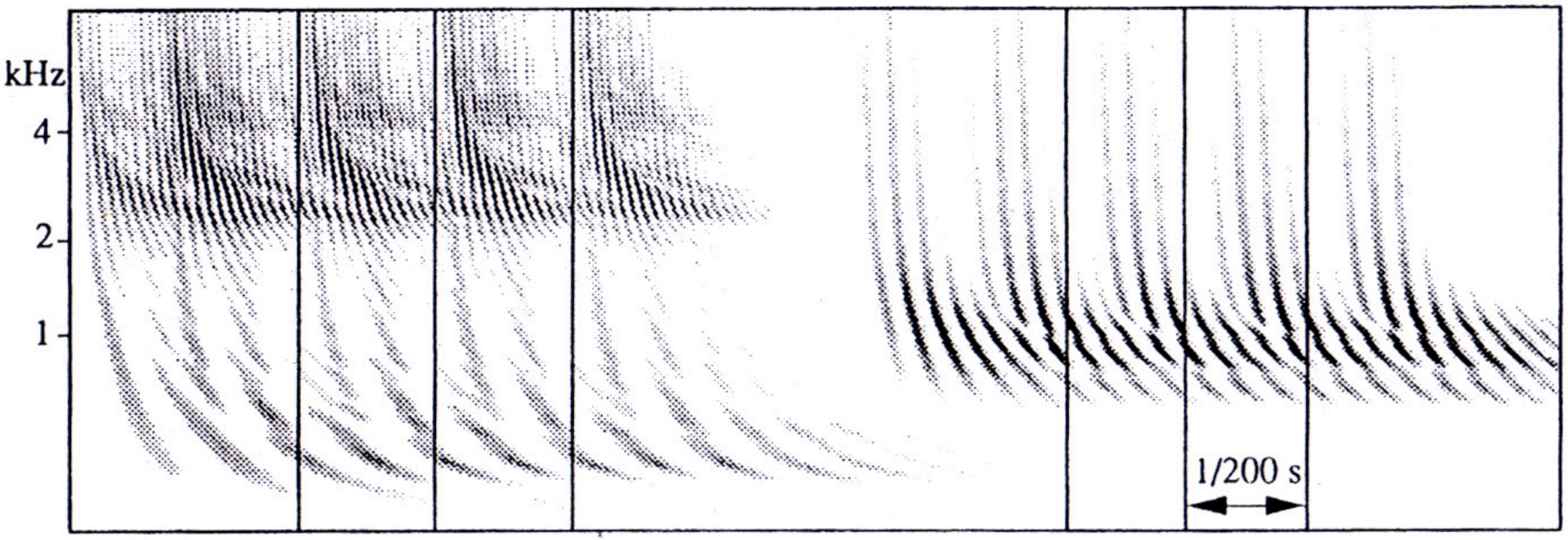
Cochlear representations



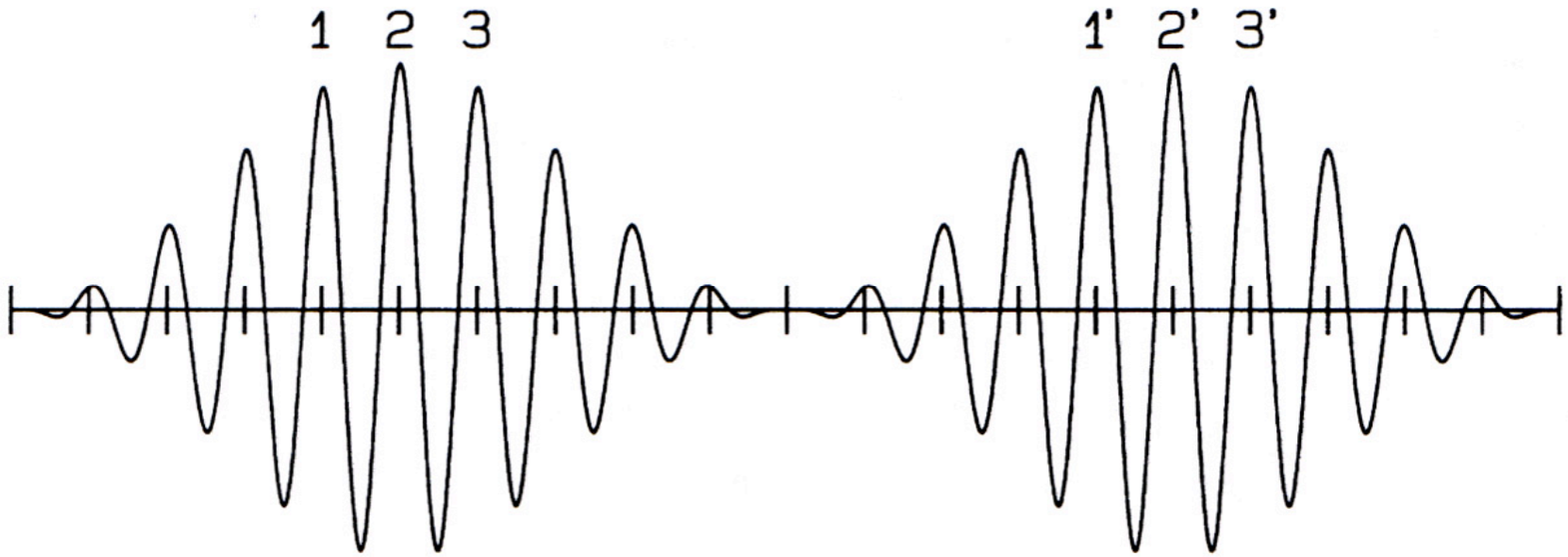
Cochlear representations

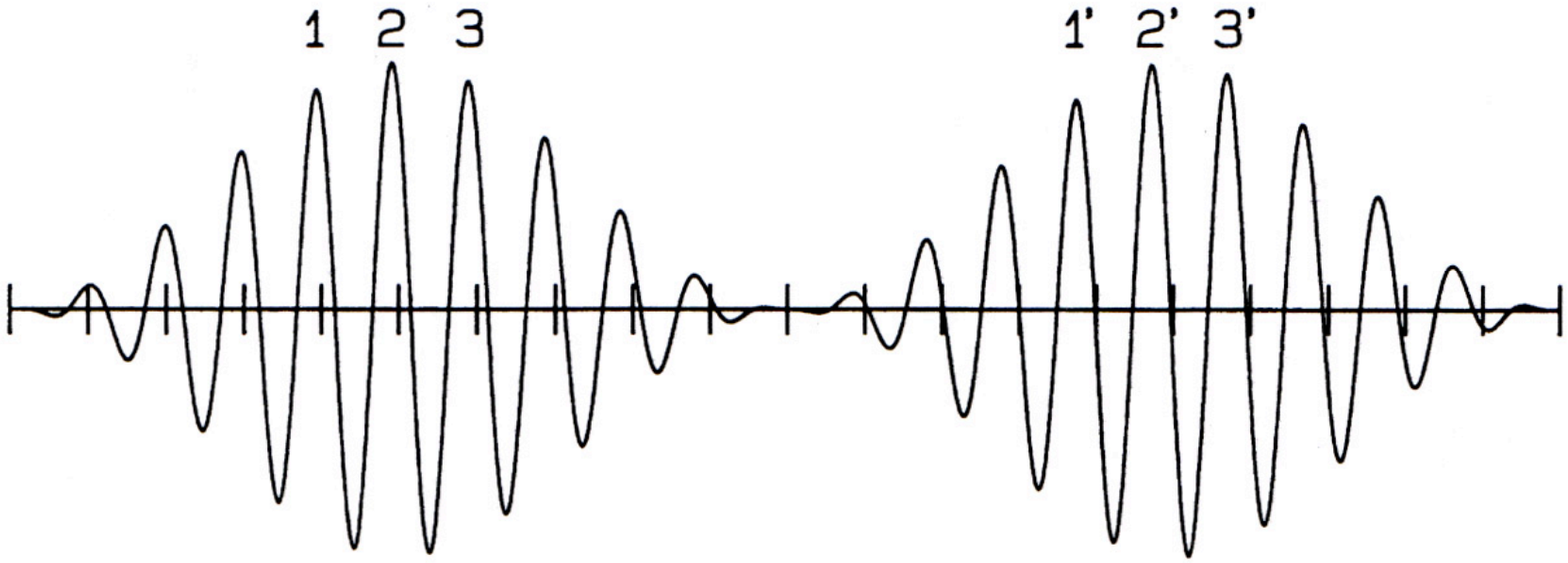


Examples of simulated cochlear output

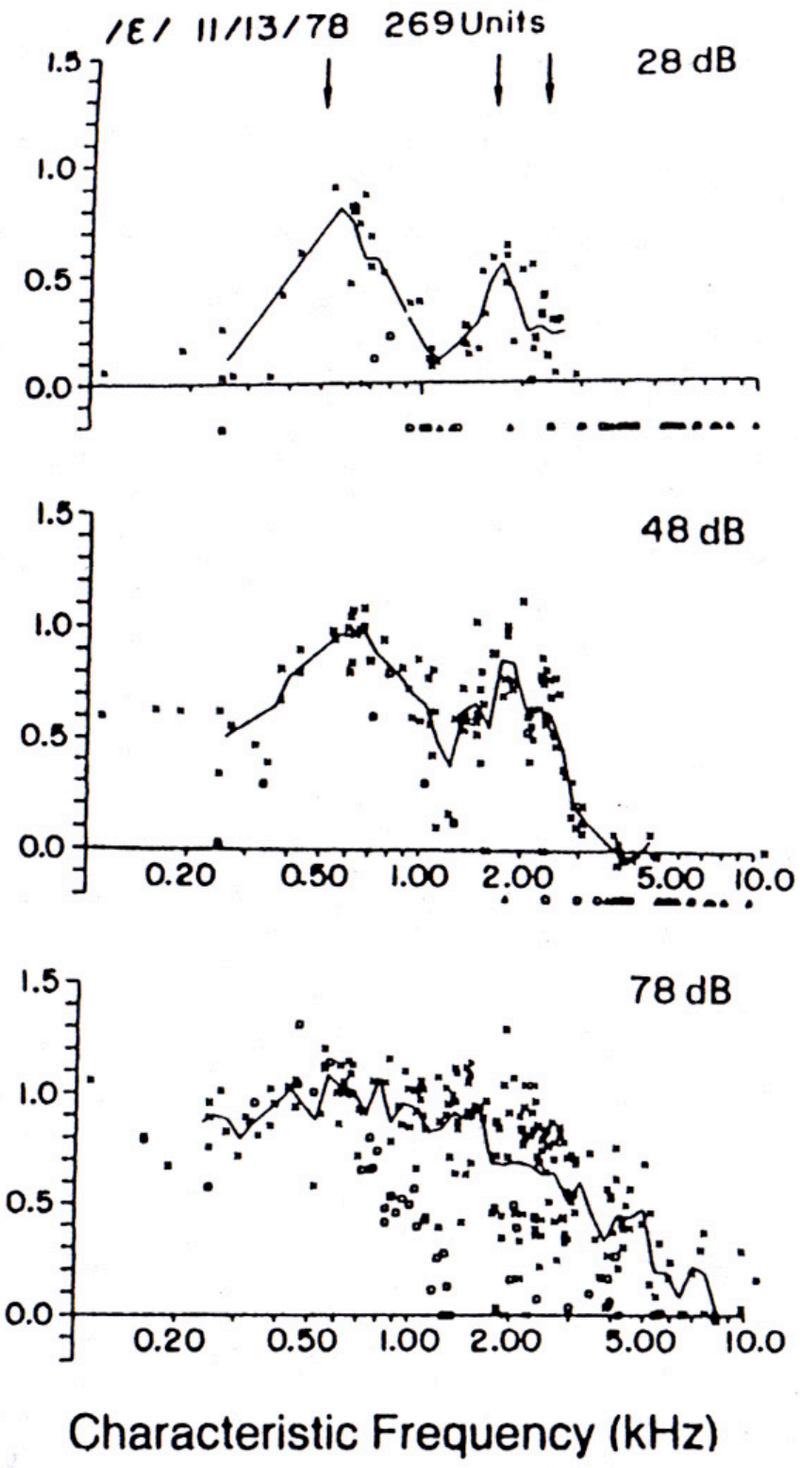


Explaining perception of multiple pitches



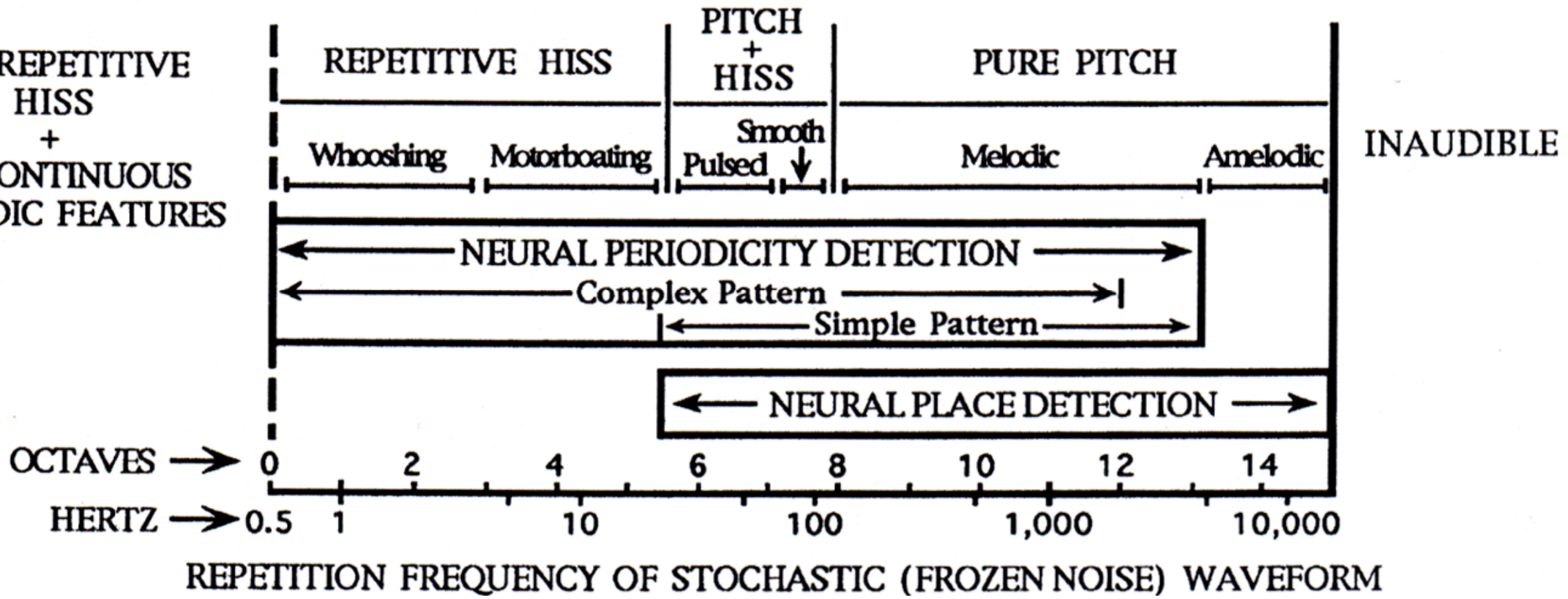


Limits of harmonic resolution

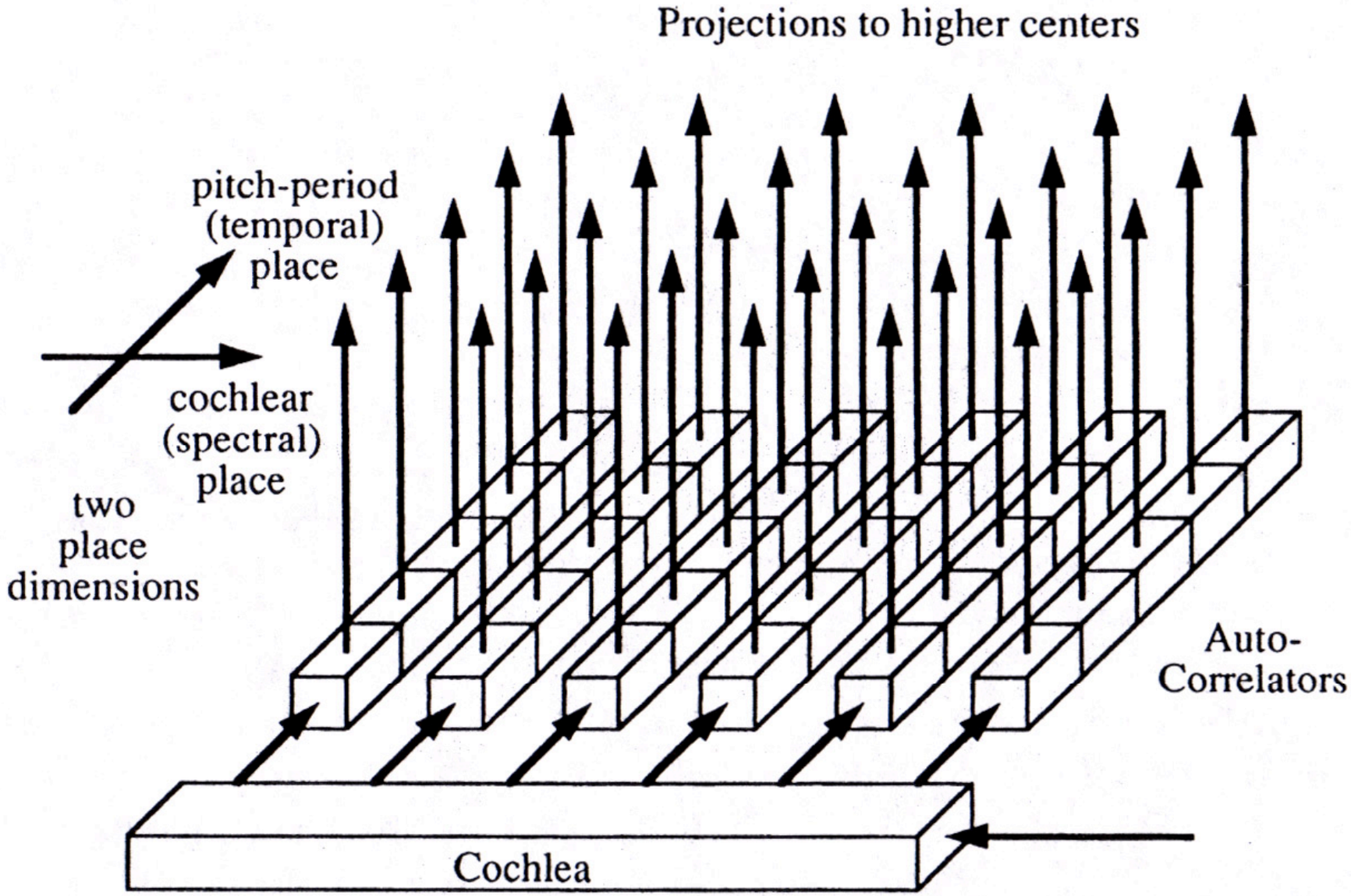


- Cochlear filters are narrow enough to resolve partially at least up to the fourth or fifth harmonic of a complex sound
- Spectral profiles computed by non-spatial algorithms can resolve up to 30 harmonics

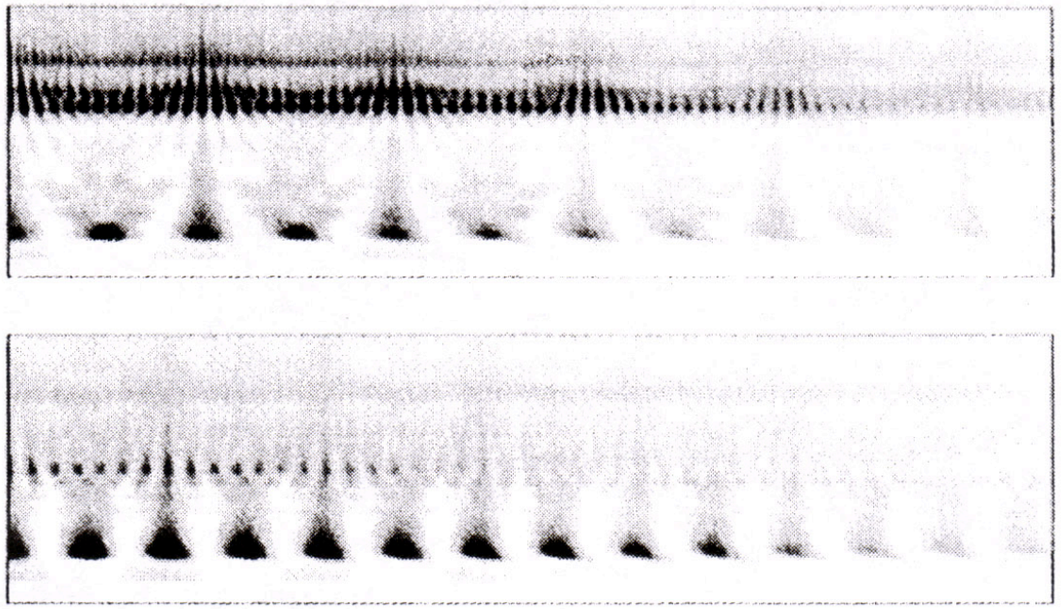
Map of pitch space



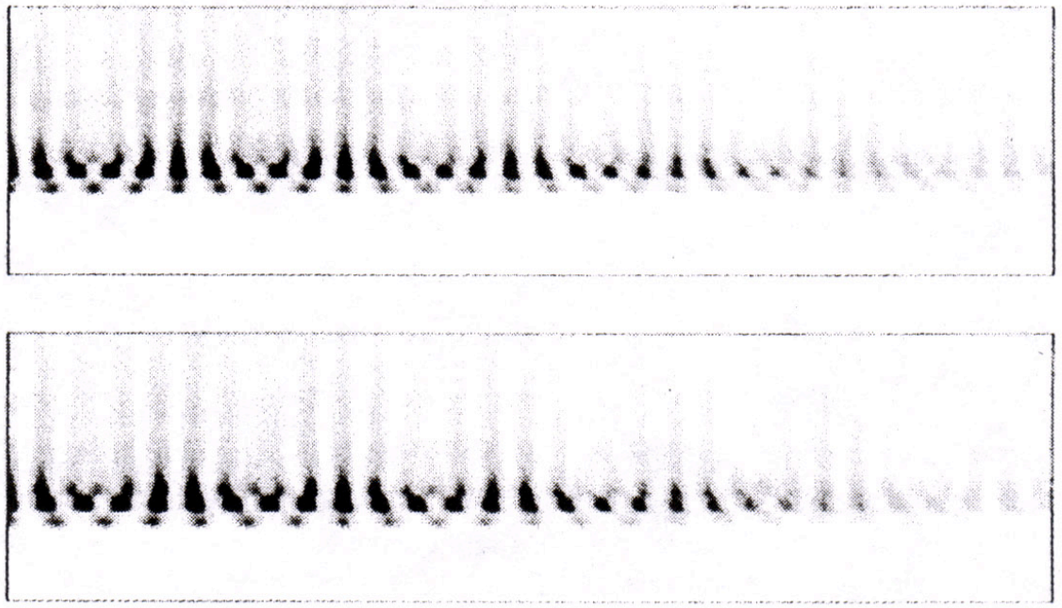
Calculating periodicity: the correlogram



Correlograms

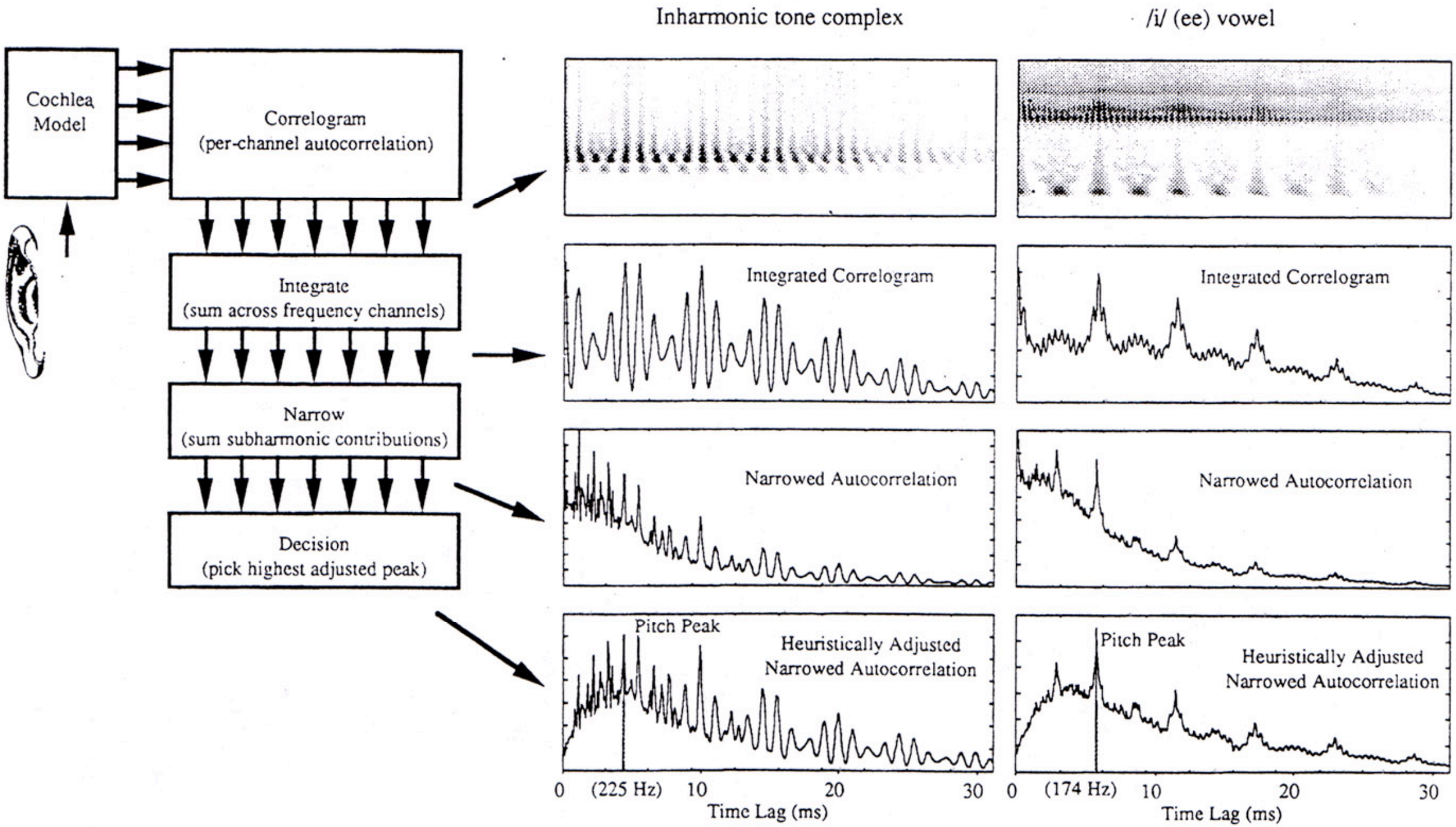


- /i/ vowel [ee]
- /u/ vowel [oo]

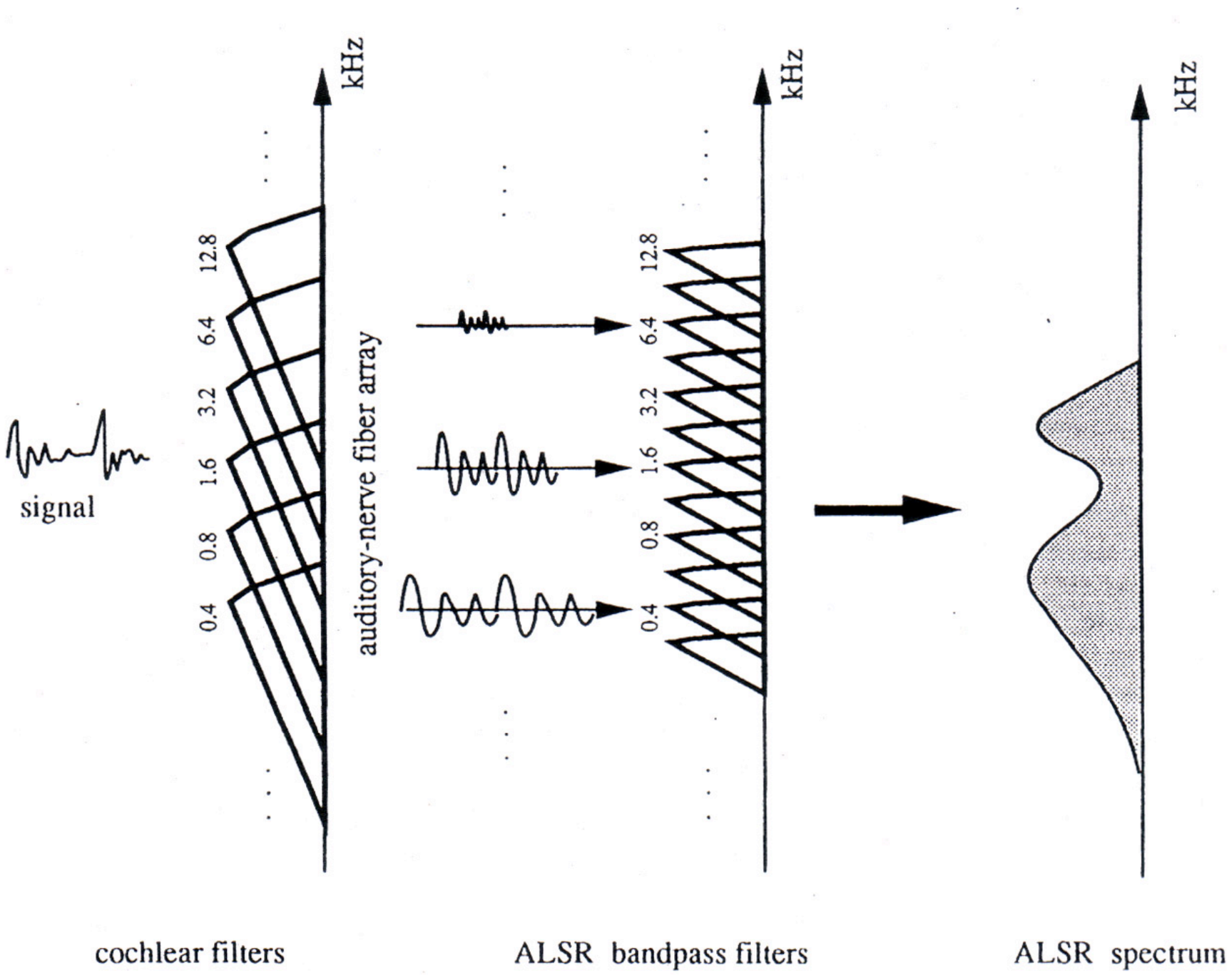


- three tone harmonic complex
- three tone inharmonic complex

Slaney-Lyon pitch model



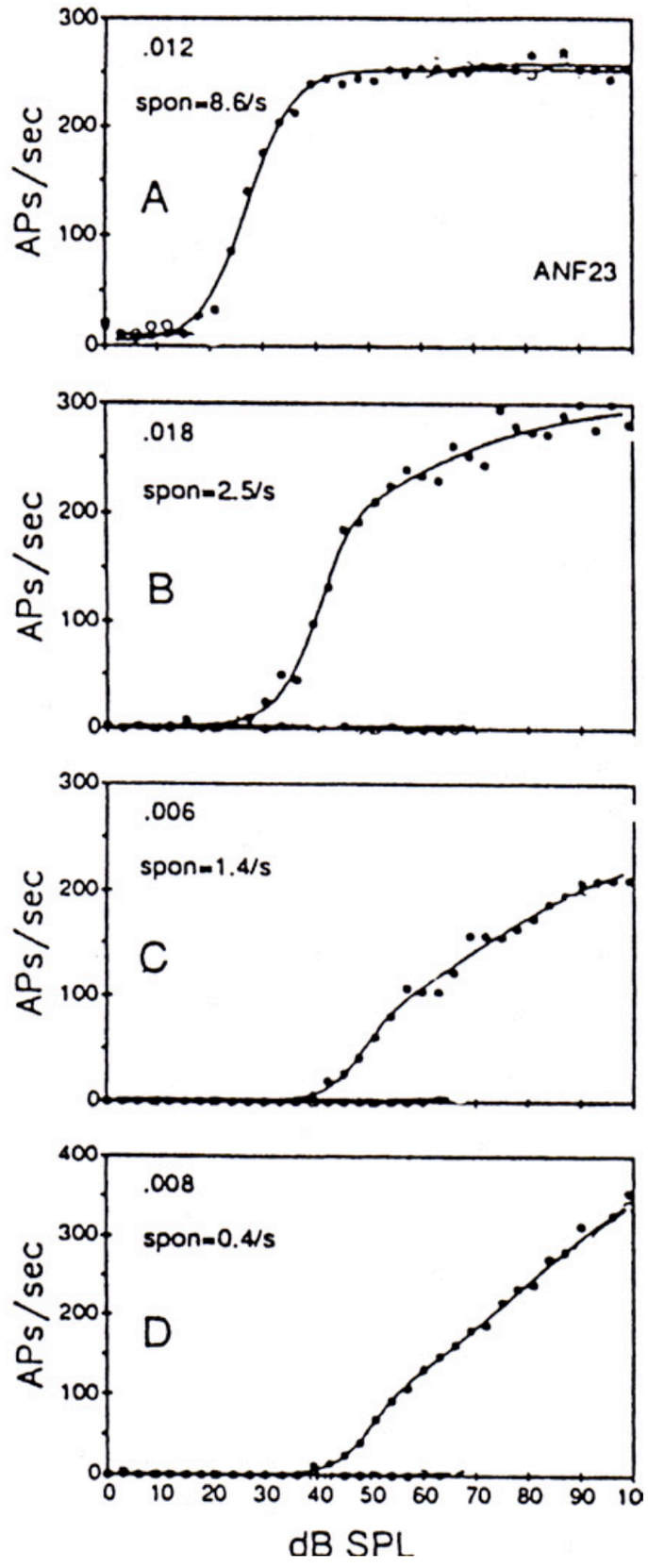
Spectral representation of timbre



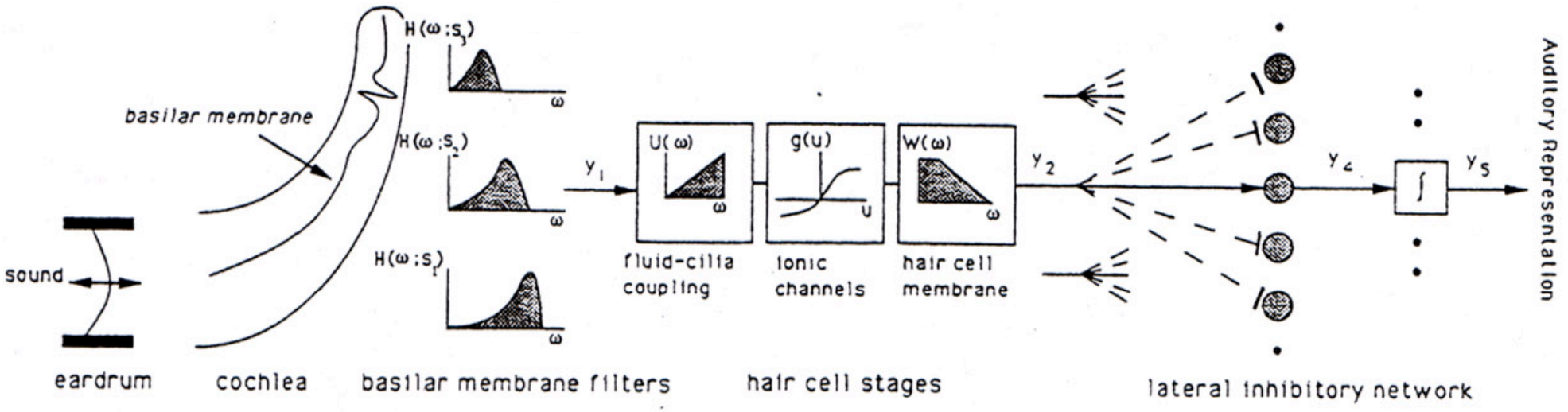
Average localized synchronous rate (Young and Sachs, 1979)

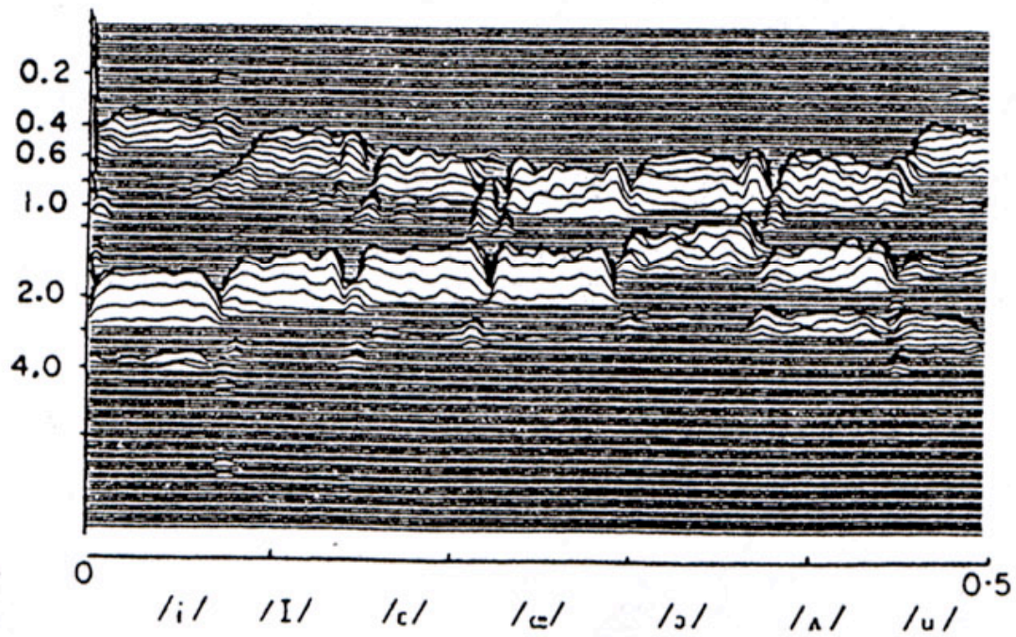
Saturation of nerve fibers

- How could the model account for saturating responses when harmonics cannot be resolved?

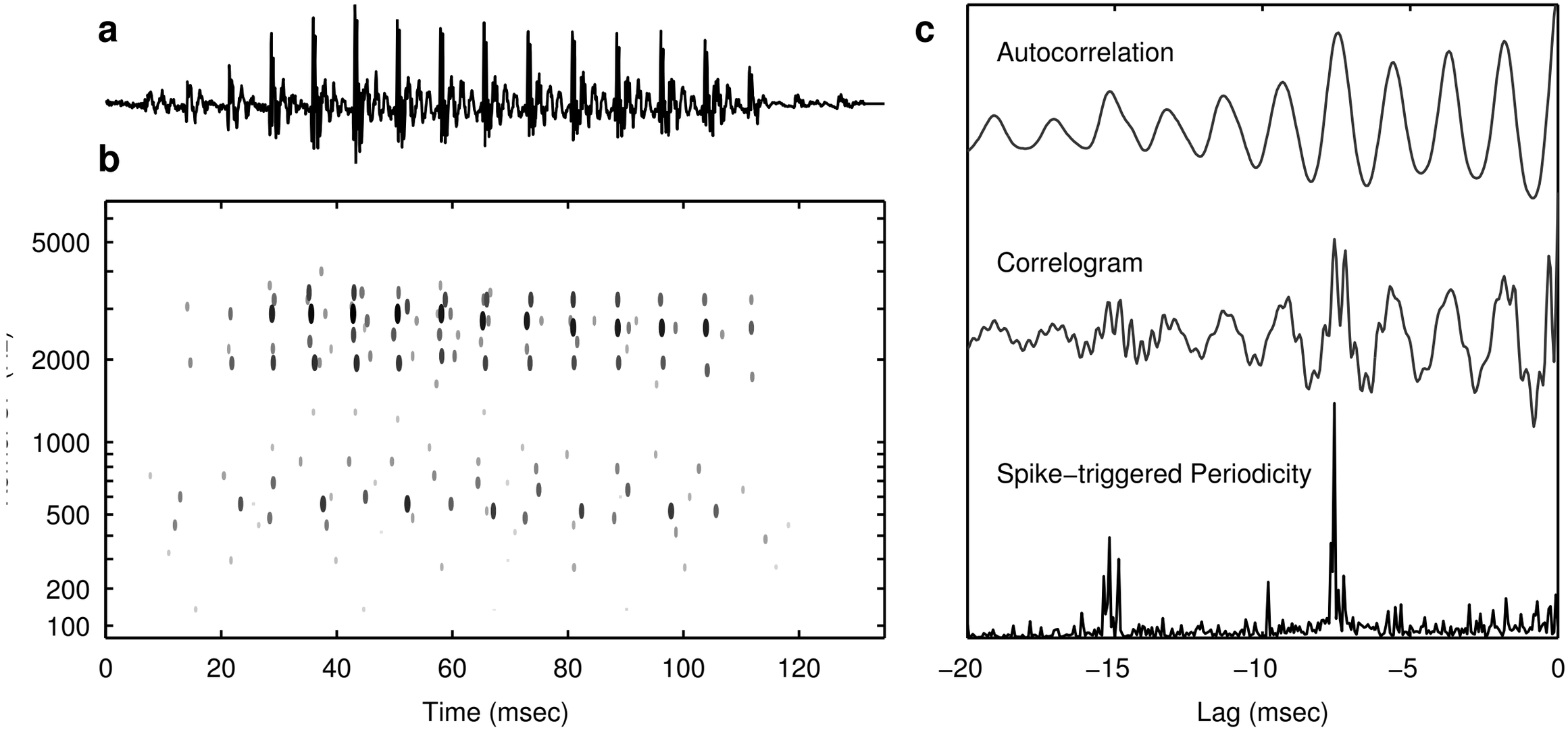


Lateral inhibitory network

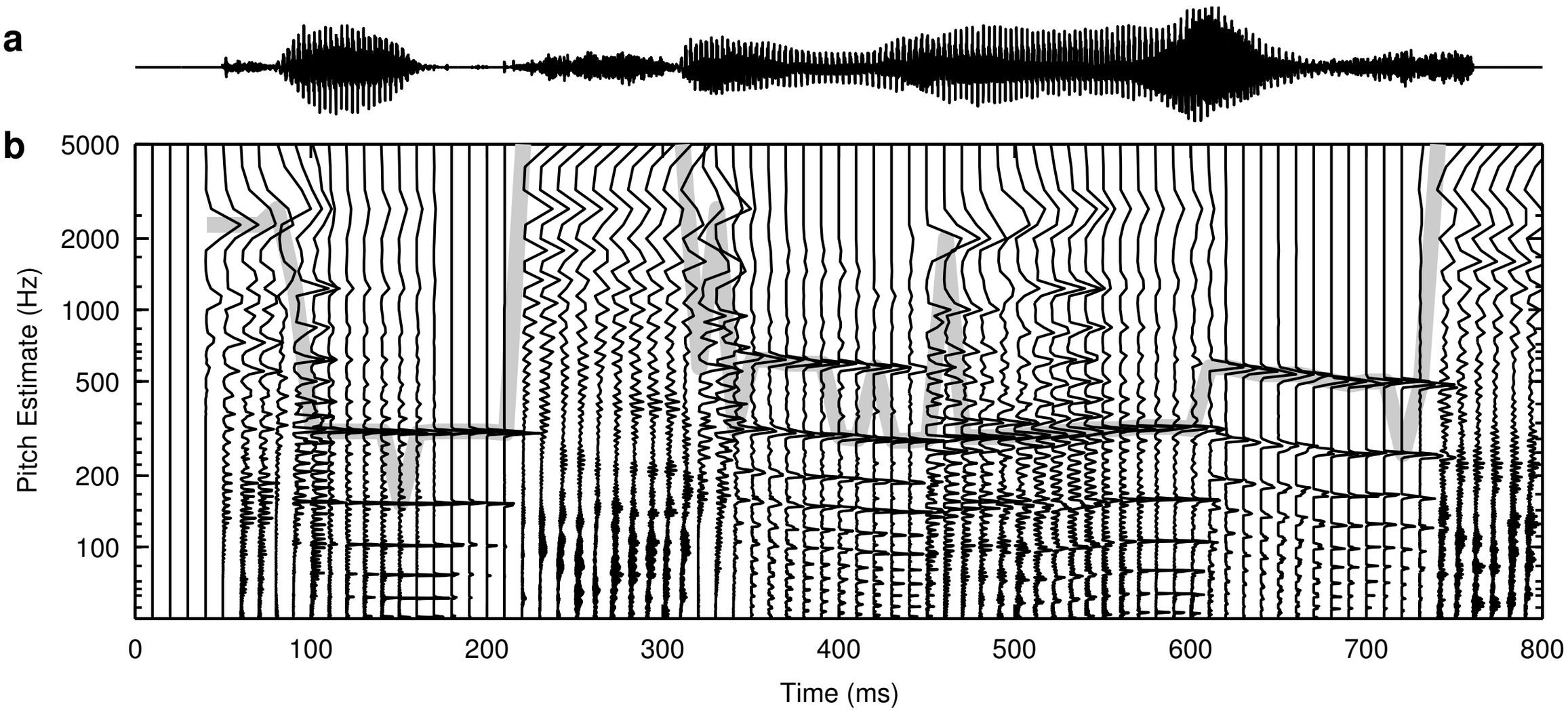




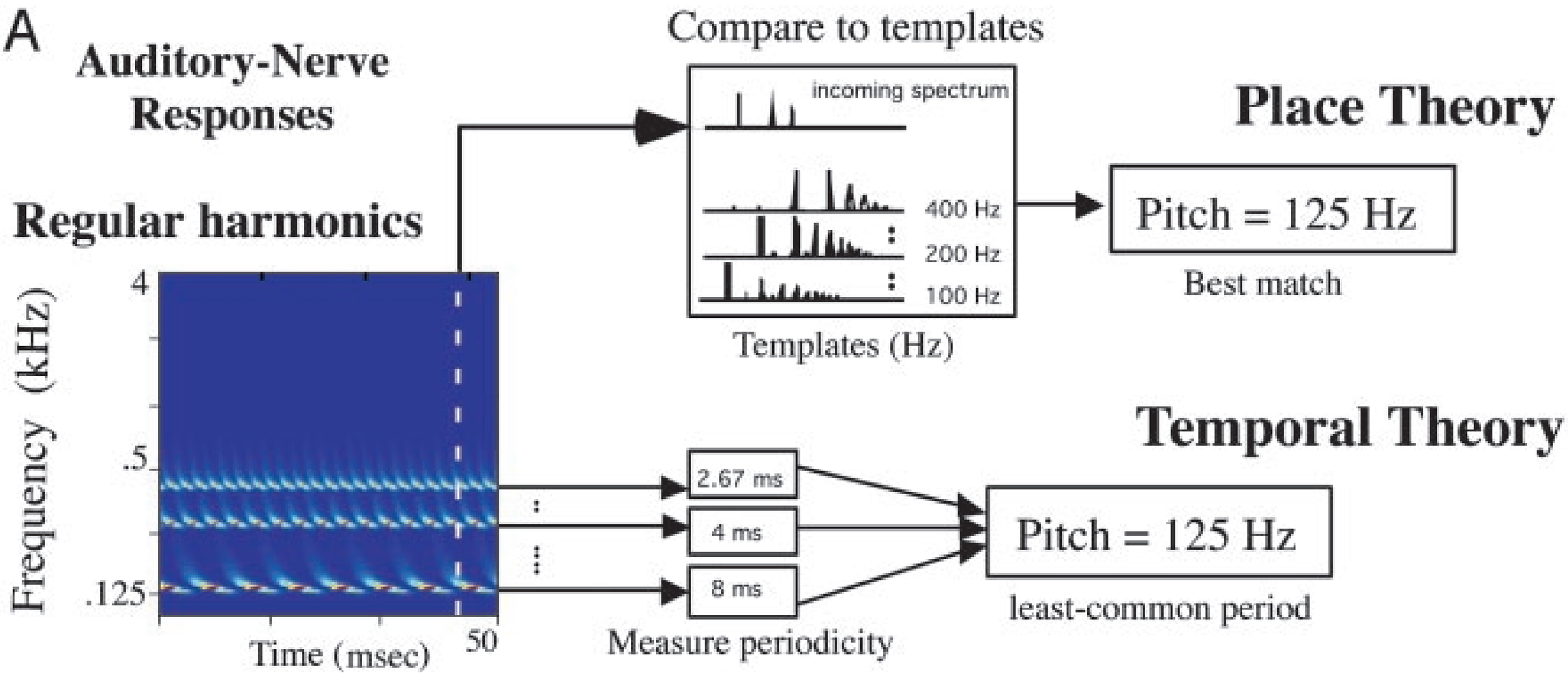
Estimating pitch with an idealized spike code



Tracking pitch over time

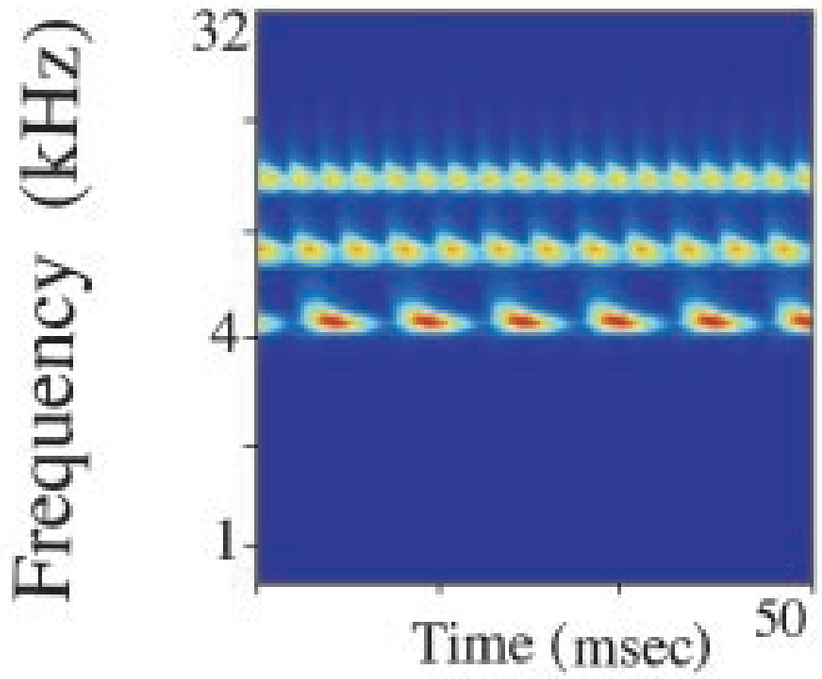


Temporal vs place theory of pitch



from Shamma (2004) on Oxenham and Bernstein (2004)

Transposed harmonics

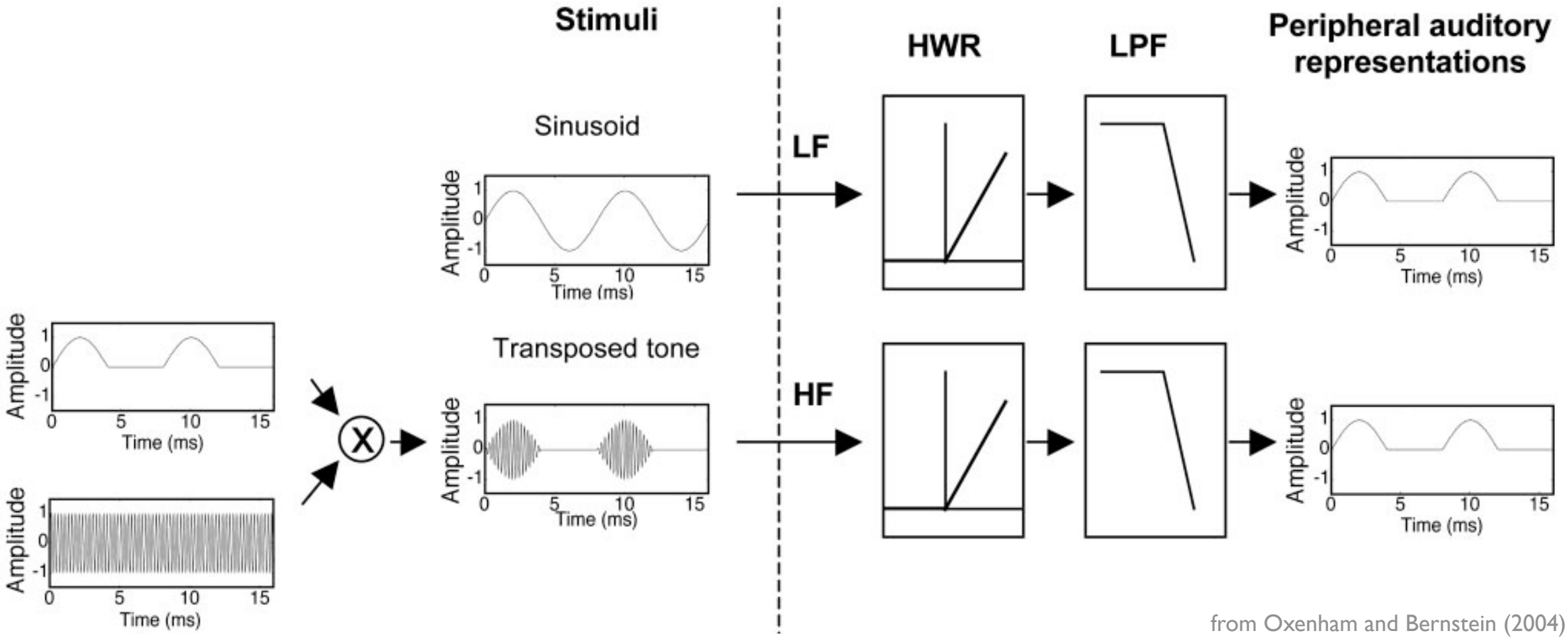


Place Theory → **NO** Pitch

Temporal Theory → Pitch = 125 Hz

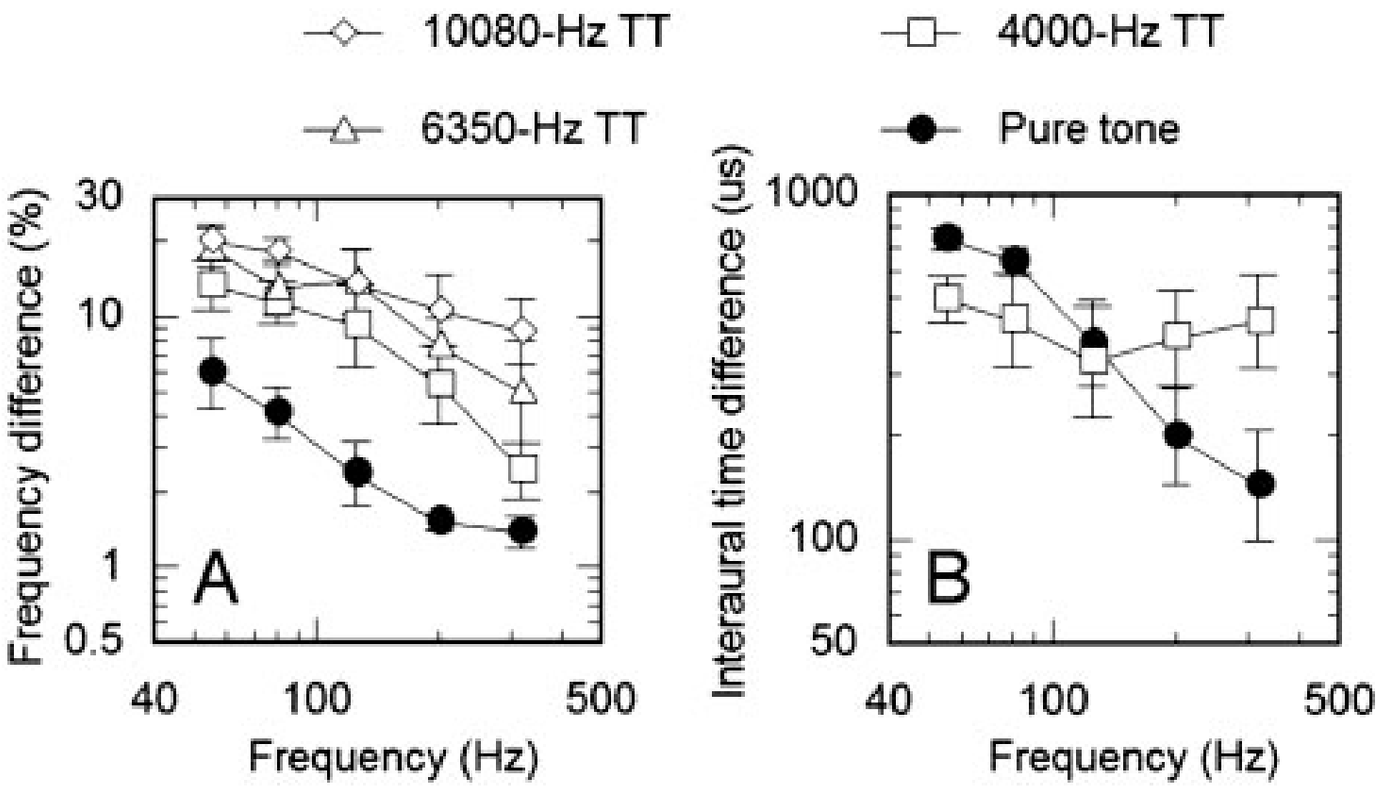
from Shamma (2004) on Oxenham and Bernstein (2004)

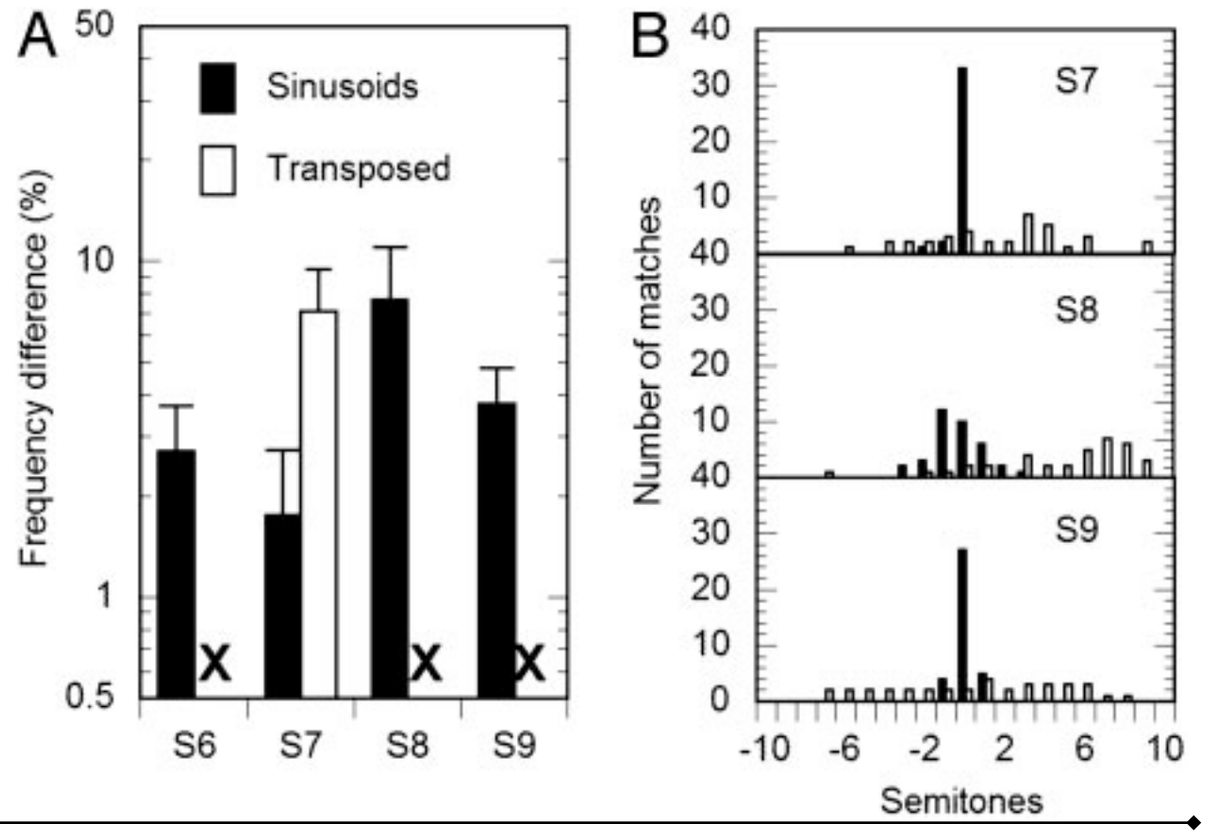
Transposing stimuli to have same temporal pitch



from Oxenham and Bernstein (2004)

Subjects show frequency but not pitch discrimination for HF tones





from Oxenham and Bernstein (2004)