Pitch Perception

HST.723. Neural Coding and Perception of Sound

Pitch Perception of Pure Tones

The pitch of a pure tone is strongly related to the tone's frequency, although there are small effects of level and masking.

<1000 Hz: increased level: decreased pitch

1000-2000 Hz: little or no change

>2000 Hz: increased level: increased pitch

Figure removed due to copyright reasons.

(Moore, 1997)

Difference Limens for Frequency (DLF)

The auditory system is exquisitely sensitive to changes in frequency (e.g. 2-3 Hz at 1000 Hz = 0.01 dB).

How is frequency coded - Place or timing?

Figure removed due to copyright reasons.

Zwicker's proposal for FM detection.
(From Moore, 1997)

- ***Place***
- Pros: Could in principle be used at all frequencies.
- Cons: Peak of BM traveling wave shifts basally with level by ½ octave – no similar pitch shift is seen; fails to account for poorer performance in DLFs at very high frequencies (> 4 kHz), although does a reasonable job of predicting frequency-modulation difference limens (FMDLs).

Temporal cues

Figure removed due to copyright reasons.

From Rose et al. (1971)

Timing

Pros: Pitch estimate is basically level-invariant; may explain the absence of musical pitch above ca. 4-5 kHz.

Cons: Thought to break down totally above about 4 kHz (although some "optimal detector" models predict residual performance up to 8 or 10 kHz); harder to explain diplacusis (differences in pitch perception between the ears).

Musical pitch

Figure removed due to copyright reasons.

Musical pitch is probably at least 2-dimensional:

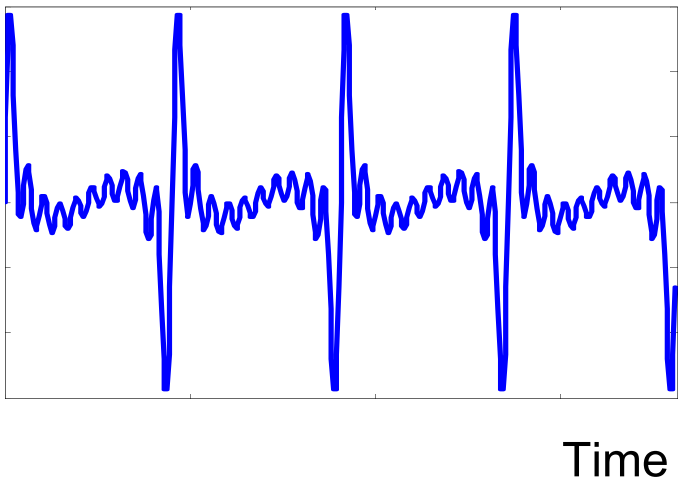
- *Tone height:* monotonically related to frequency
- *Tone chroma:* related to pitch class (note name)

Circularity in pitch judgments: changes in chroma but no change in height. In circular pitch is a half-octave interval perceived as going up or down? (Deutsch, 1987)

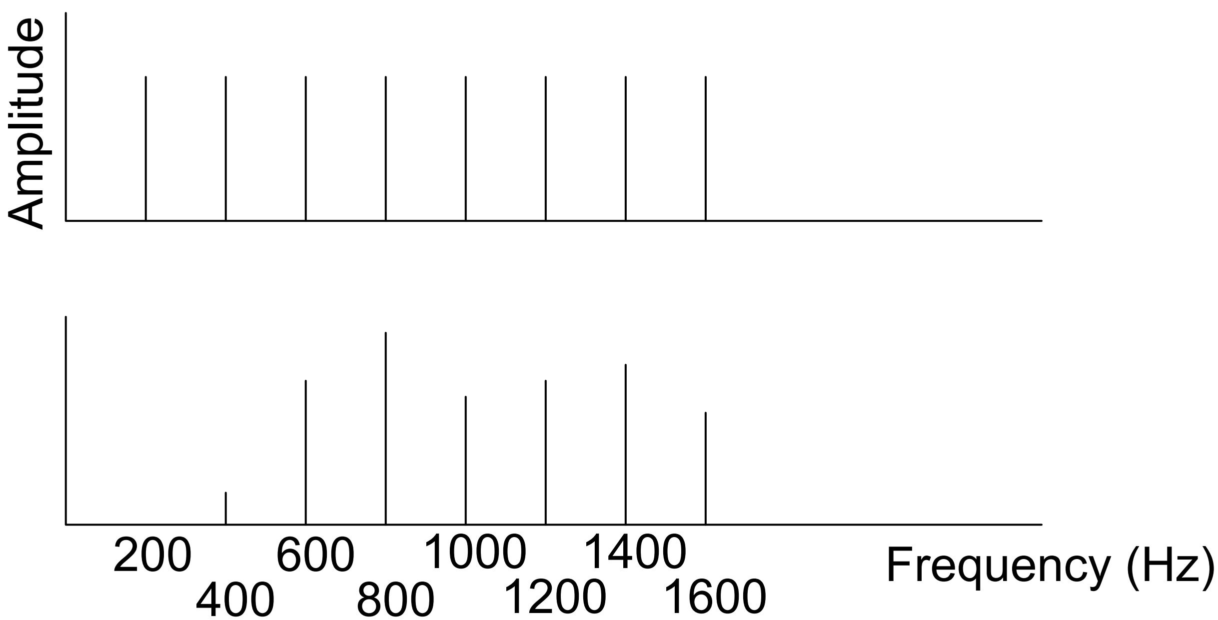
- Musical pitch of pure tones breaks down above about 5 kHz: octave matches become erratic and melodies are no longer recognized. Differences in frequency are still detected – only tone chroma is absent.
- Further evidence for the influence of temporal coding?

(Demo from ASA Auditory Demonstrations CD)

Pitch of complex tones



- Harmonic tones produce a pitch at the fundamental frequency (F0), even if there is no energy at the F0 itself *(pitch of the missing fundamental)*. Evidence against Ohm/Helmholtz place theory.



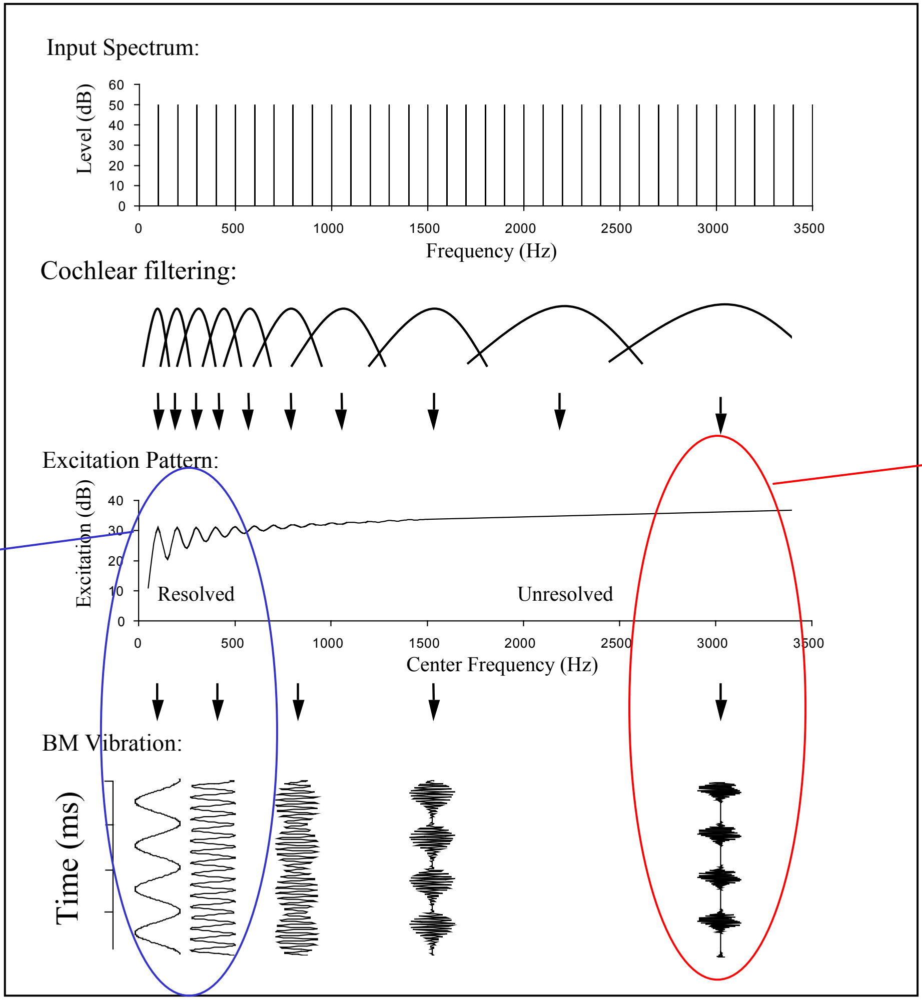
Pitch = 200 Hz

Pitch = 200 Hz

Harmonic complex tones

Many sounds in our world are harmonic complex tones, consisting of many sinusoids all at multiples of the *fundamental frequency* (F0).

Resolved harmonics: Temporal fine structure

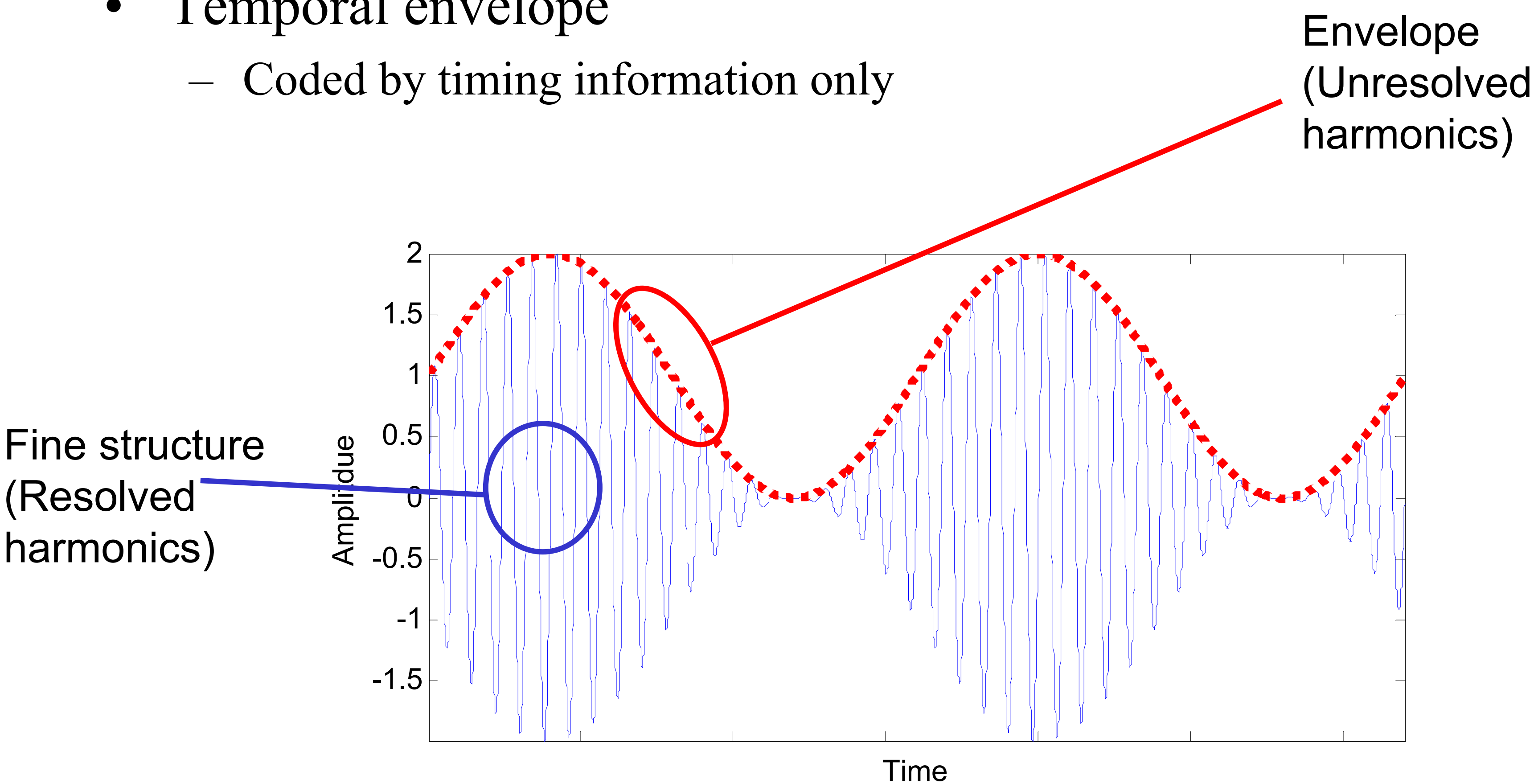


Unresolved harmonics: Temporal envelope

(Plack & Oxenham, 2005)

Two temporal cues in complex sounds

- Temporal fine structure
  - Could be coded either by place or time (or both)
- Temporal envelope
  - Coded by timing information only



High (unresolved) harmonics produce poor musical pitch

Unresolved	Highpass filtered above 8^{th} harmonic
Resolved	Lowpass filtered below 8^{th} harmonic
Resolved & Unresolved	No filtering

(Courtesy of Bertrand Delgutte.)

Low (resolved) harmonics dominate pitch perception

100-100 100-106 100-112 100-133 100-178

F0 below 800 Hz F0 above 800 Hz

Figure removed due
to copyright reasons.

Resynthesized sentences with low- and high-spectral regions on different F0s (Demo by C.J. Darwin)

Mechanisms of Complex Pitch Perception: The Early Years

Temporal Theory (Schouten, 1940):
Pitch is extracted from the summed waveform of adjacent components. This requires that some components interact.

Pattern Recognition Theory (e.g. Goldstein, 1973):
The frequencies of individual components are determined and the "best-fitting" f0 is selected. This requires that some components remain resolved *and* that some form of "harmonic template" exists.

Pros and Cons of Temporal and Place Models of Pitch

Evidence against a "pure" temporal model

- Pitch sensation is strongest for low-order (resolved) harmonics (Plomp, 1967; Ritsma, 1967).
- Pitch can be elicited by only two components, one in each ear (Houtsma and Goldstein, 1972).
- Pitch can be elicited by consecutively presented harmonics (Grose et al., 2002).

Evidence again a "pure" pattern recognition theory

- Very high, unresolved harmonics can still produce a (weaker) pitch sensation
- Aperiodic, sinusoidally amplitude-modulated (SAM) white noise can produce a pitch sensation (Burns and Viemeister, 1976; 1981).

Autocorrelation model of pitch perception

Figure removed due to copyright considerations.
Please see: Meddis, R., and M. Hewitt.
“Virtual pitch and phase sensitivity studied
of a computer model of the auditory periphery.
I: Pitch identification.” *J Acoust Soc Am 89*
(1991): 2866-2882.

- Based on an original proposal by Licklider (1951).
- The stimulus within each frequency channel is correlated (delayed, multiplied and averaged) with itself (through delay lines).
- This produces peaks at time intervals corresponding to multiples of the stimulus period.
- Pooling interval histograms across frequency produces an overall estimate of the “dominant” interval, which generally corresponds to the fundamental frequency.

Autocorrelation model

Pros:

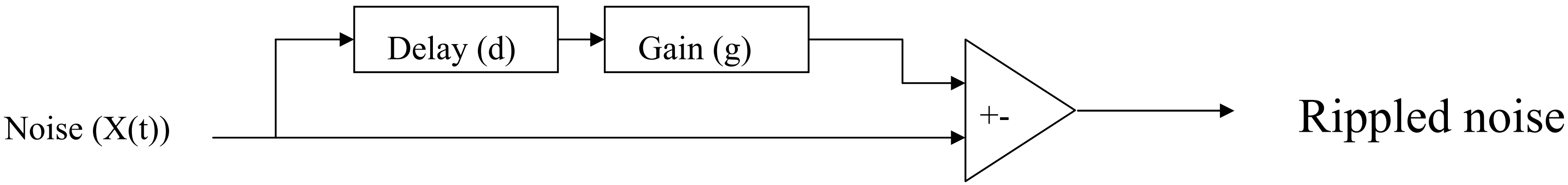
- Model can deal with both resolved and unresolved harmonics
- Predicts no effect of phase for resolved harmonics, but strong phase effects for unresolved harmonics, in line with data (Meddis & Hewitt, 1991).
- Predicts a dominance region of pitch, roughly in line with early psychophysical data, due to reduction in phase locking with frequency.

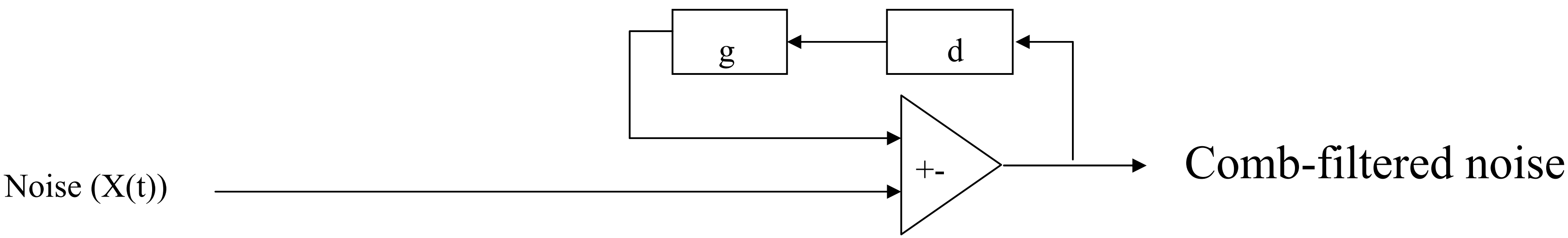
Cons:

- Deals *too well* with unresolved harmonics – predicts no difference based on resolvability, in contrast to psychophysical data (Carlyon and Shackleton, 1994).
- Dominance region based on absolute, not relative, frequency, in contrast to data.

[N.B. The "template" model of Shamma and Klein (2000) involves place and timing coding, but not in the traditional sense.]

“Regular Interval Noise”





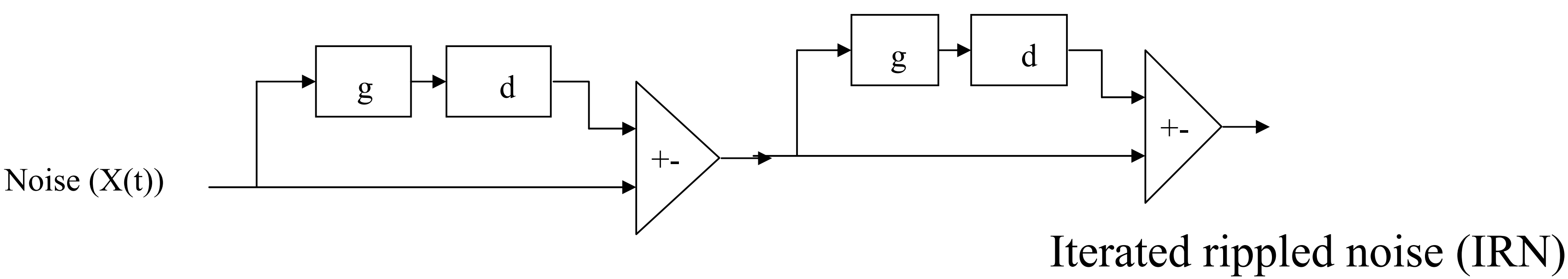


Figure removed due
to copyright reasons.

Patterson et al. (2002)

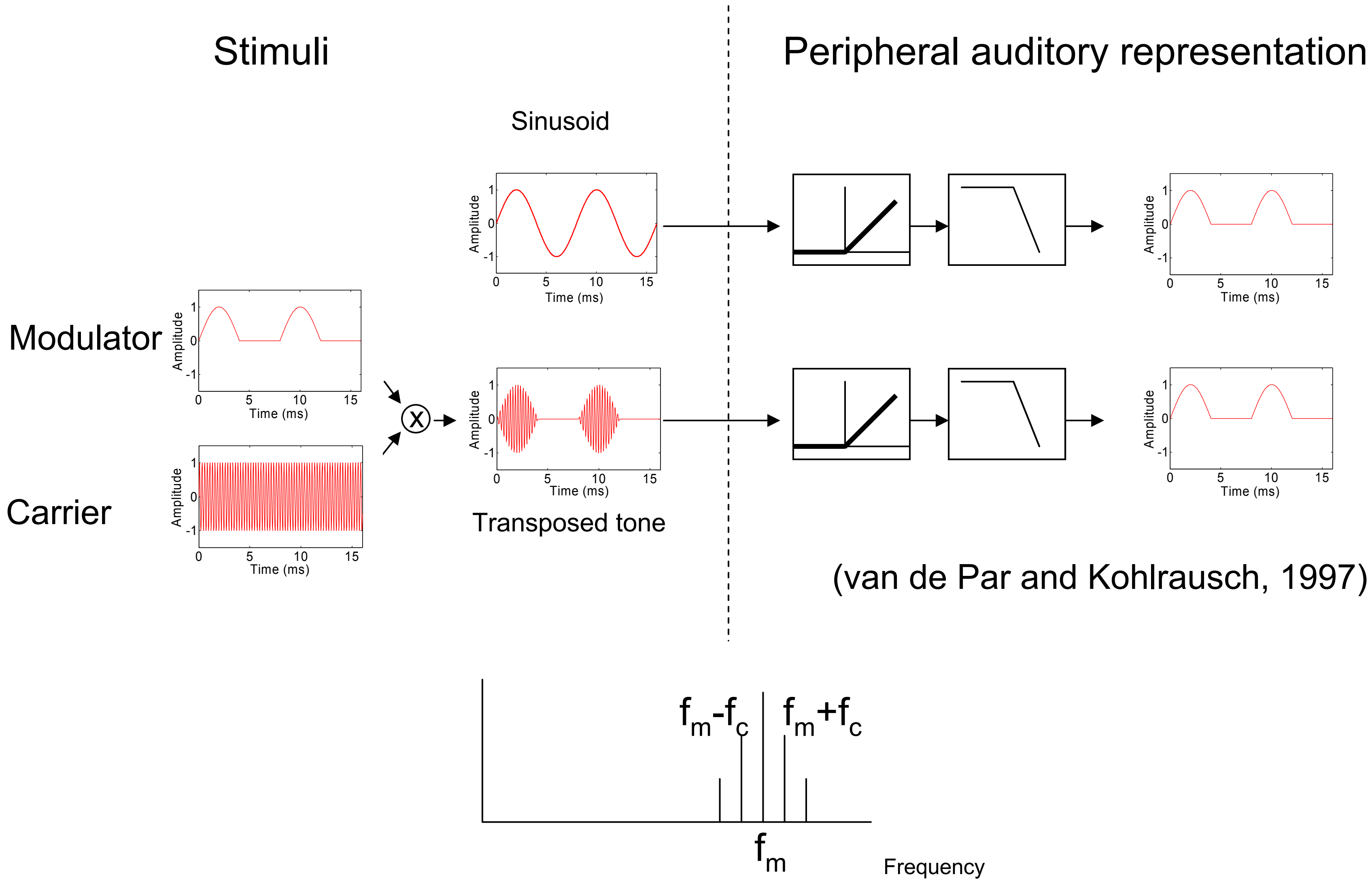
Distinguishing time from place

- For pure tones, temporal and place information co-vary, making dissociation difficult.
- *Transposed stimuli* (van de Par & Kohlrausch, 1997) are an attempt to overcome this.

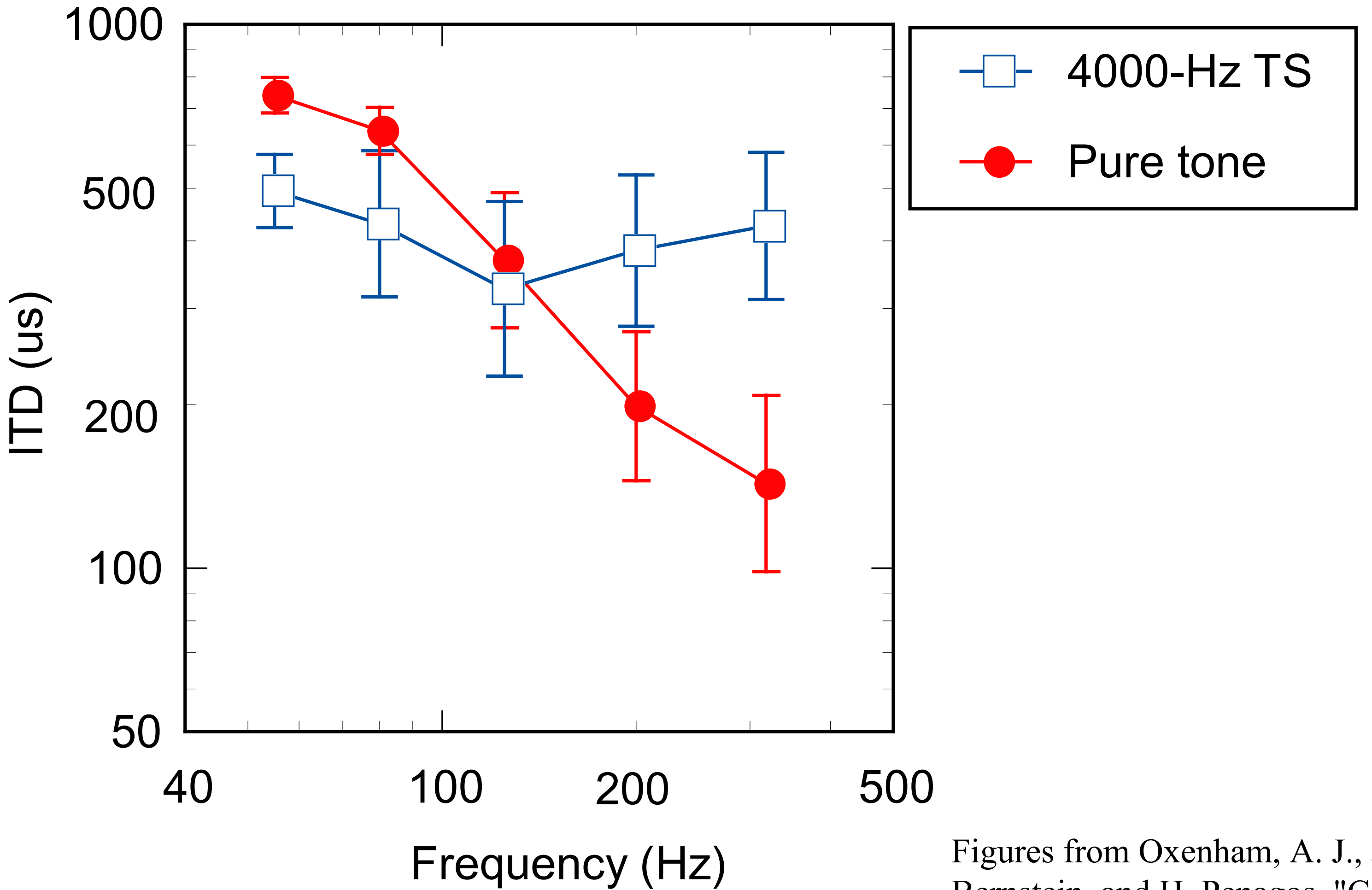
AIMS:

- Transpose low-frequency temporal *fine-structure* information into the *envelope* of a high-frequency carrier.
- Dissociate place and time representations.

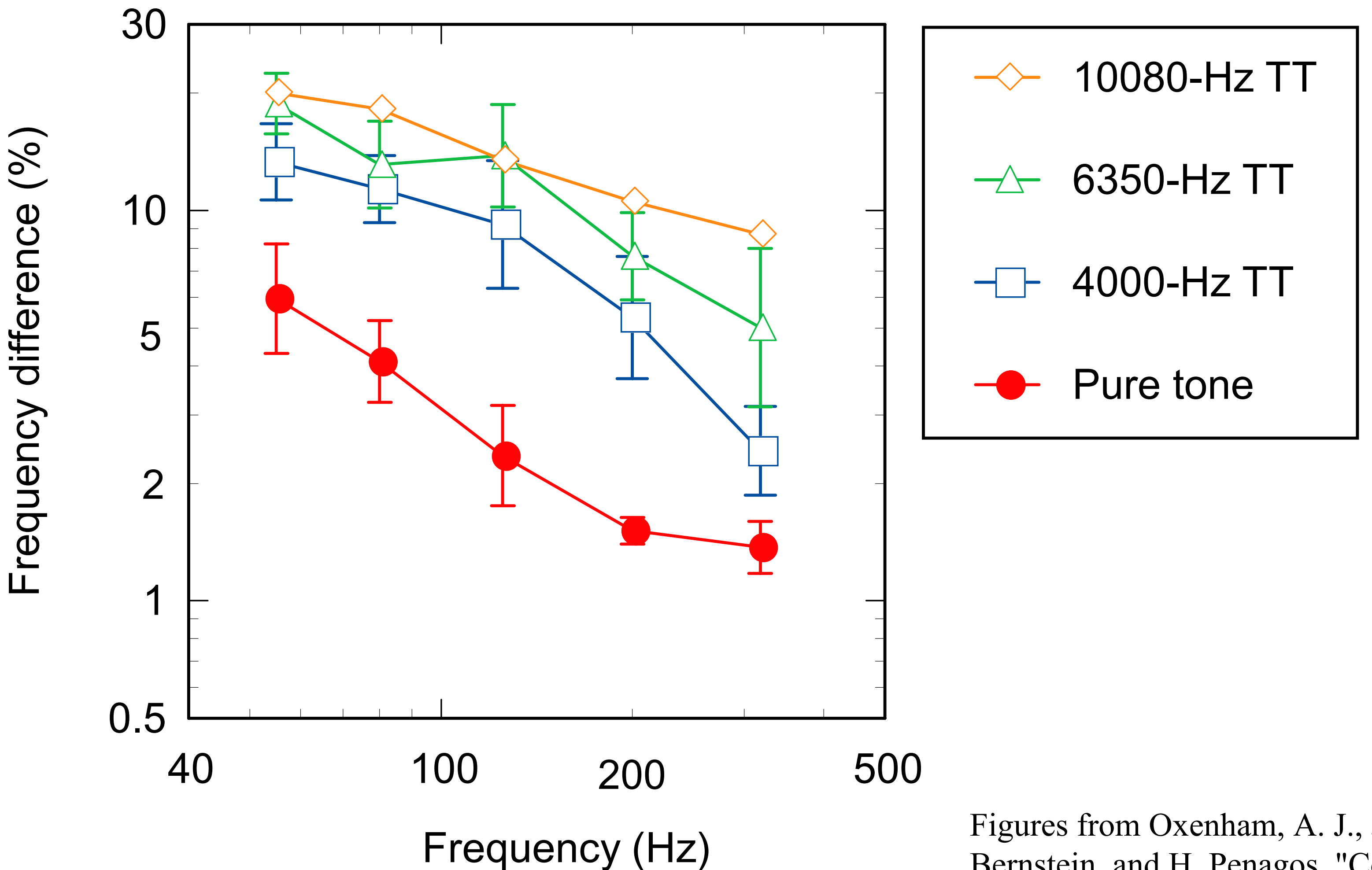
What are transposed stimuli?



Interaural Time Differences (ITDs)



Pure-tone frequency difference limens



Figures from Oxenham, A. J., J. G. W. Bernstein, and H. Penagos. "Correct tonotopic representation is necessary for complex pitch perception," *Proc Natl Acad Sci USA* 101 (2004): 1421-1425.

Transposed tones: Simple pitch

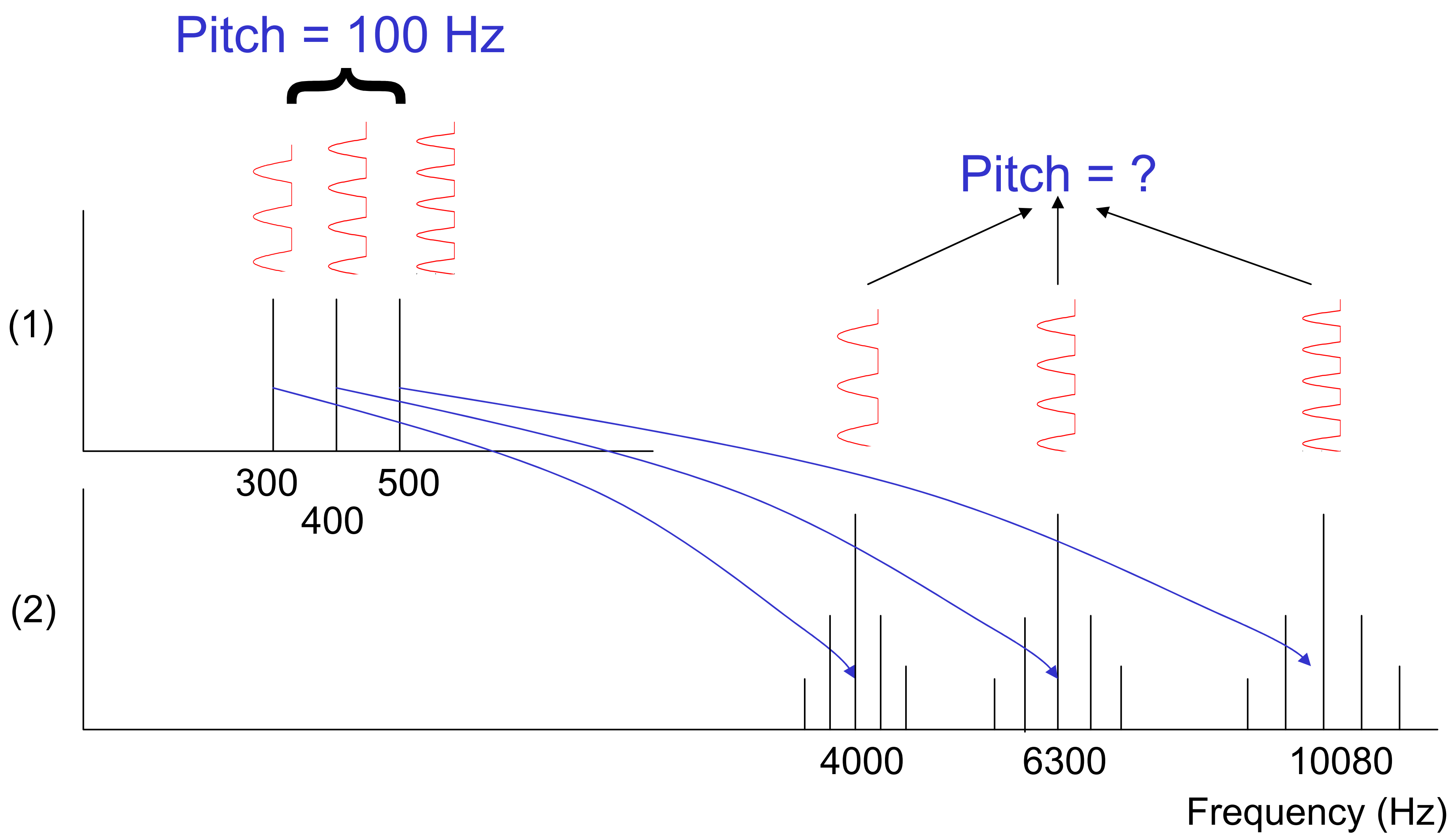
- Unlike ITDs, temporal information for frequency cannot be used optimally by the auditory system.
- Pitch perception seems weaker for all transposed tones.
- Place information may be important.

300-Hz pure tone

300-Hz tone, transposed to 4 kHz

What about complex pitch?

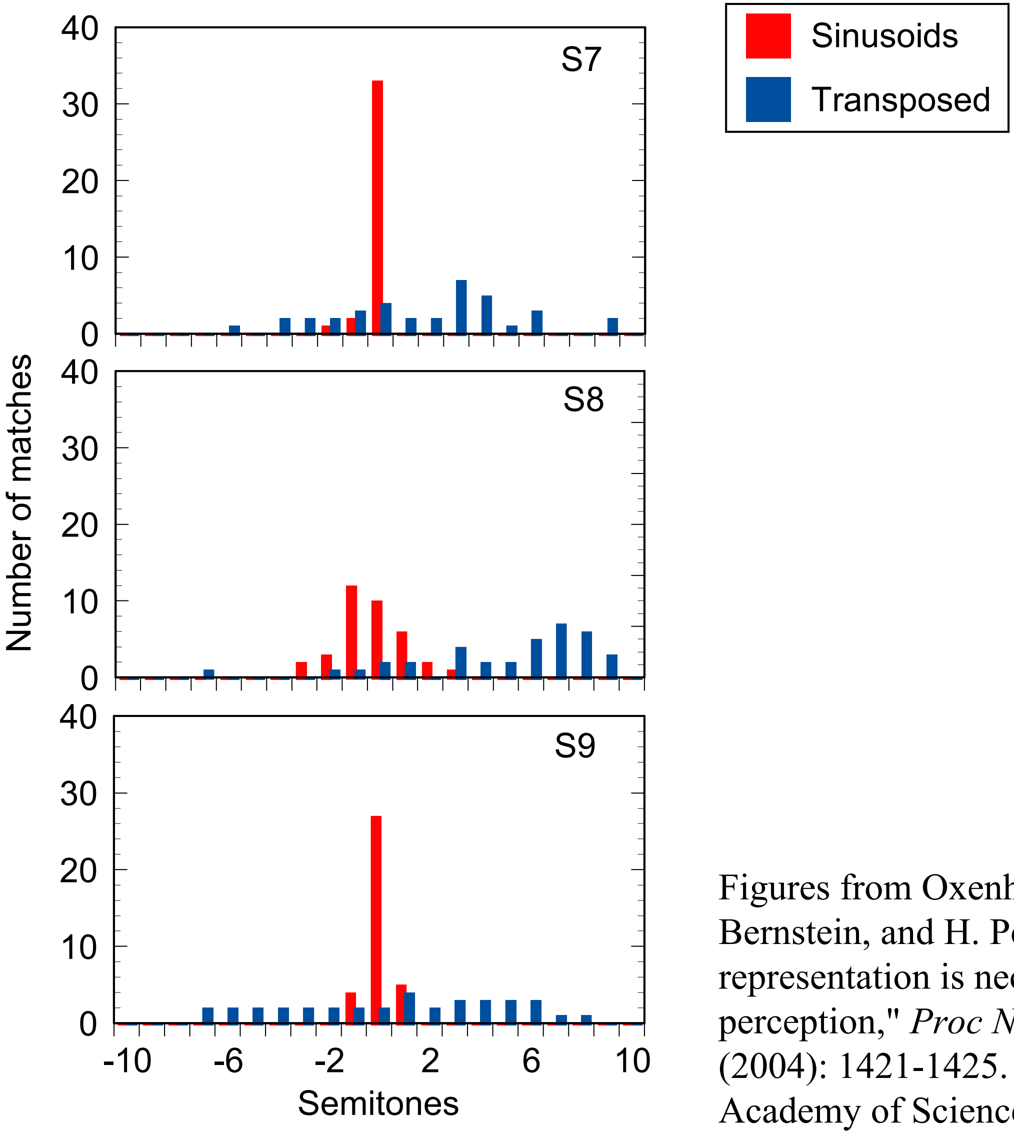
Complex tone pitch perception



Temporal model predictions

Figure removed due to copyright considerations.
Please see: Meddis, R., and L. O'Mard. "A unitary model of pitch perception." *J Acoust Soc Am* 102 (1997): 1811-1820.

Pitch matches



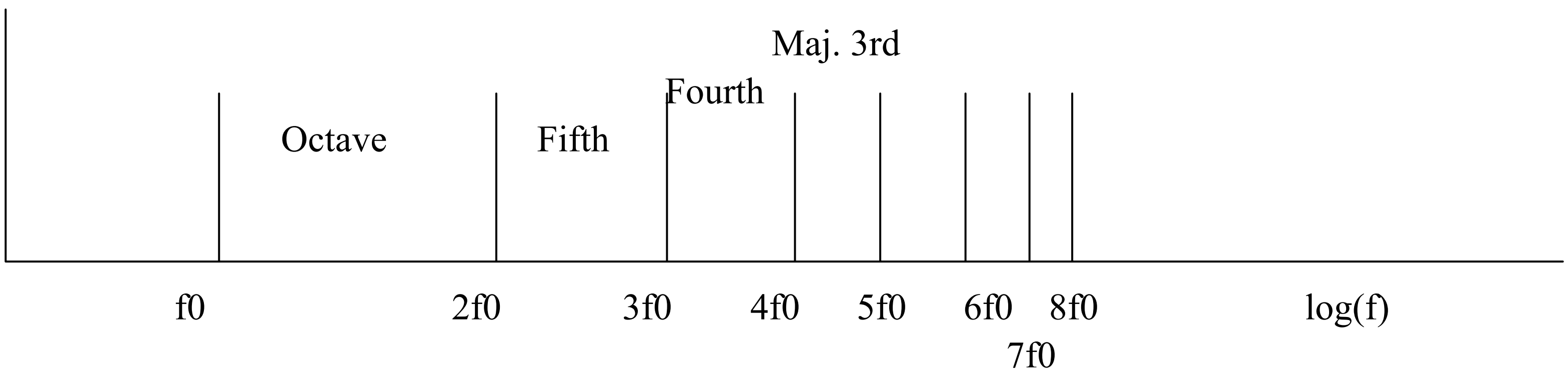
Figures from Oxenham, A. J., J. G. W. Bernstein, and H. Penagos. "Correct tonotopic representation is necessary for complex pitch perception," *Proc Natl Acad Sci USA* 101 (2004): 1421-1425. Copyright (2004) National Academy of Sciences, U.S.A.

Transposed tones: Conclusions

- Pitch of pure tones is poor and complex pitch is nonexistent.
- Suggests that fine structure must be presented to the correct *place* in the cochlea – timing is not enough.
- Possible hybrid models include Shamma et al.'s (2000) harmonic template model.

Musical intervals: Consonance and Dissonance

- In the West, the equal- (or well-) tempered scale has been adopted, with the octave split into twelve equal (semitone) steps on a log scale, i.e., 1 semitone higher is $2^{1/12}$ times higher in frequency.
- This is a compromise: the intervals in the harmonic series only approximate the notes of the scale.



- Perceived dissonance is in part due to beating effects between neighboring harmonics. Remaining effect of perceived consonance and dissonance may be simply cultural.

Auditory Grouping and Pitch

Simultaneous, harmonically related tones tend to form a single auditory object, which makes ecological sense.

What happens if one component is slightly out of tune?

Harmonicity can be a strong cue in binding components together, but it can be overridden by competing cues or expectations (Darwin et al., 1994; 1995).

A mistuned harmonic can be "heard out" more easily, but can still contribute to the overall pitch of the complex. This is an example of "duplex perception".

References

Moore, B. C. J. (**1997**). An Introduction to the Psychology of Hearing (Academic Press, London).

Rose, J. E., Hind, J. E., Anderson, D. J., and Brugge, J. F. (**1971**). "Some effects of the stimulus intensity on response of auditory nerve fibers in the squirrel monkey," J. Neurophysiol. **34**, 685-699.

Deutsch, D. (**1987**). "The tritone paradox: effects of spectral variables," Percept Psychophys **41**, 563-575.

Plack, C. J., and Oxenham, A. J. (**2005**). "Pitch perception," in Pitch: Neural Coding and Perception, edited by C. J. Plack, A. J. Oxenham, A. N. Popper and R. Fay (Springer, New York).

Schouten, J. F. (**1940**). "The residue and the mechanism of hearing," Proc. Kon. Akad. Wetenschap. **43**, 991-999.

Goldstein, J. L. (**1973**). "An optimum processor theory for the central formation of the pitch of complex tones," J. Acoust. Soc. Am. **54**, 1496-1516.

Ritsma, R. J. (**1967**). "Frequencies dominant in the perception of the pitch of complex sounds," J. Acoust. Soc. Am. **42**, 191-198.

Plomp, R. (**1967**). "Pitch of complex tones," J. Acoust. Soc. Am. **41**, 1526-1533.

Houtsma, A. J. M., and Goldstein, J. L. (**1972**). "The central origin of the pitch of complex tones: Evidence from musical interval recognition," J. Acoust. Soc. Am. **51**, 520-529.

Grose, J. H., Hall, J. W., and Buss, E. (**2002**). "Virtual pitch integration for asynchronous harmonics," J. Acoust. Soc. Am. **112**, 2956-2961.

Burns, E. M., and Viemeister, N. F. (**1976**). "Nonspectral pitch," J. Acoust. Soc. Am. **60**, 863-869.

Burns, E. M., and Viemeister, N. F. (**1981**). "Played again SAM: Further observations on the pitch of amplitude-modulated noise," J. Acoust. Soc. Am. **70**, 1655-1660.

Licklider, J. C. R. (**1951**). "A duplex theory of pitch perception," Experientia **7**, 128-133.

Meddis, R., and Hewitt, M. (**1991**). "Virtual pitch and phase sensitivity studied of a computer model of the auditory periphery. I: Pitch identification," J. Acoust. Soc. Am. **89**, 2866-2882.

Shamma, S., and Klein, D. (**2000**). "The case of the missing pitch templates: How harmonic templates emerge in the early auditory system," J. Acoust. Soc. Am. **107**, 2631-2644.

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., and Griffiths, T. D. (**2002**). "The processing of temporal pitch and melody information in auditory cortex," Neuron **36**, 767-776.

van de Par, S., and Kohlrausch, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," J. Acoust. Soc. Am. **101**, 1671-1680.

Oxenham, A. J., Bernstein, J. G. W., and Penagos, H. (**2004**). "Correct tonotopic representation is necessary for complex pitch perception," Proc. Natl. Acad. Sci. USA **101**, 1421-1425.

Meddis, R., and O'Mard, L. (**1997**). "A unitary model of pitch perception," J. Acoust. Soc. Am. **102**, 1811-1820.

Darwin, C. J., Ciocca, V., and Sandell, G. J. (**1994**). "Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic.," Journal of the Acoustical Society of America **95**, 2631-2636.

Darwin, C. J., Hukin, R. W., and al-Khatib, B. Y. (**1995**). "Grouping in pitch perception: Evidence for sequential constraints," J. Acoust. Soc. Am. **98**, 880-885.