Perception of pitch

AUDL4007: 11 Feb 2010. A. Faulkner.

See Moore, BCJ "Introduction to the Psychology of Hearing, Chapter 5".

Or Plack CJ "The Sense of Hearing" Lawrence Erlbaum, 2005 Chapter 7

Definitions

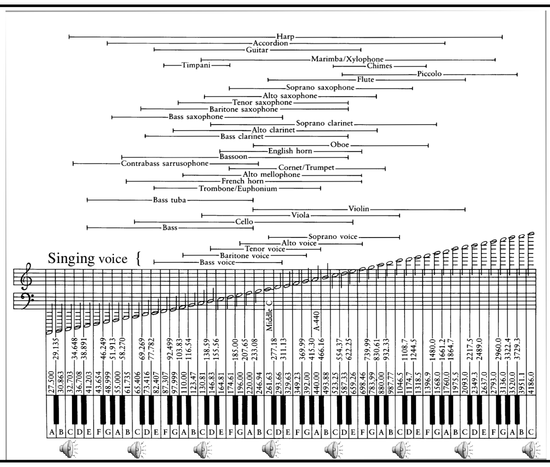
Perception: Pitch is the perceptual property of sound that conveys melody

Acoustics: Pitch is closely related to frequency and periodicity

Pitch is a perceptual property of periodic and approximately periodic sounds – these have spectra that contain harmonics of a common fundamental frequency.

Pitch should be distinguished from "timbre", which is a perceptual quality relating to the sharpness of dullness of a sound. Timbre is mainly related to spectral shape

The pitch of a sound is defined, for the purposes of measurement, as being equivalent to the frequency of a simple sine wave that has the same pitch as the sound. Hence pitch is expressed in Hz.



Why is pitch important?

- In speech
  - Pitch variations signal differences between child, adult male and adult female speakers.
  - Pitch variation conveys intonation, which indicates lexical stress and aspects of syntax.
    - e.g. *it's raining?* "checking" question usually shows final pitch rise
    - *No I mean the BLUE shirt!* – emphasis on BLUE would lead to pitch rise
  - In tone languages, pitch movement is lexically contrastive

STANDARD CHINESE **ma**

	Chinese Character	Tone symbol	Tone description
mother	媽	˥	high level
hemp	麻	˧˥	high rising
horse	馬	˨˩˦	low falling
scold	罵	˥˩	high falling

Importance of pitch: 2

- Music
- Separating sources of sound
  - Pitch is rather like a carrier frequency that we can tune in to
- Much studied in examining roles of spectral and temporal coding and processing in hearing

Auditory coding of frequency and pitch

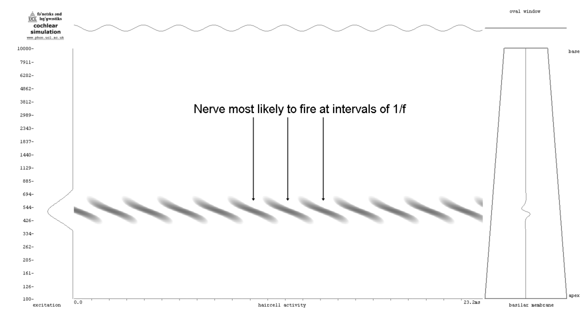
Information in spectral/place and time domains

Theories of pitch perception have been largely concerned with contrasting the contributions of spectral and temporal cues to the perception of pitch.

- Place representation - pitch is related to place of basilar membrane vibration
- Temporal representation - neural firing pattern preserves periodicity of the signal

7

Place and time coding of sine-wave frequency



Pitch Discrimination for sinewaves

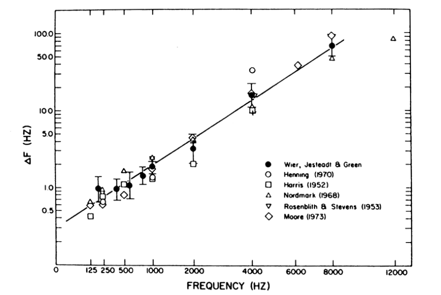


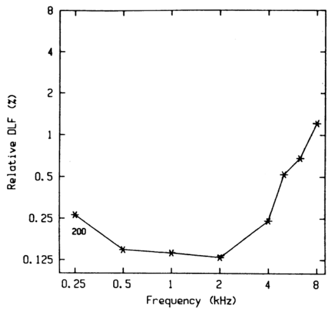
FIG. 5.1 Summary of the results of several studies measuring frequency discrimination thresholds. The thresholds, ΔF, are plotted in Hz as a function of frequency. From Wier *et al.* (1977).

Practiced listeners can hear differences of less than 1 Hz for a 200 Hz sinusoid (precision better than 0.5%)

At 1000 Hz, differences of 2 Hz can be detected (precision of about 0.2%)

NB Scales here chosen to fit data to straight line: square root(F) and threshold frequency difference on a log scale

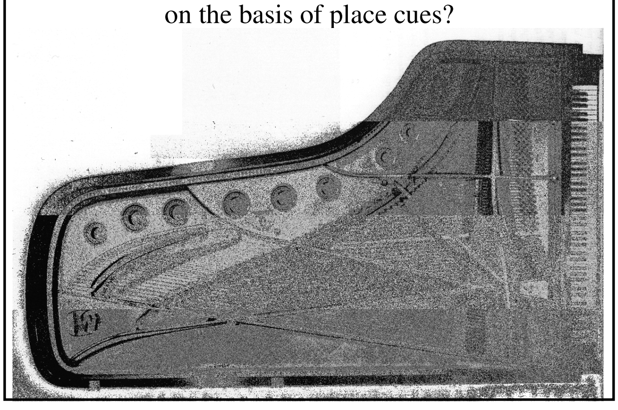
Relative discriminability of pitch



Typically pitch discrimination is expressed relative to frequency. Expressed this way the relative Difference Limen for Frequency (DLF) is smallest at 2 kHz.

10

Can we account for pure tone discrimination on the basis of place cues?



Excitation pattern coding of frequency difference

Intensity discrimination thresholds are about 1 dB.

At 1000 Hz for excitation levels to differ by 1dB requires a frequency difference of about 10 Hz – yet we can here a frequency difference of 2 Hz.

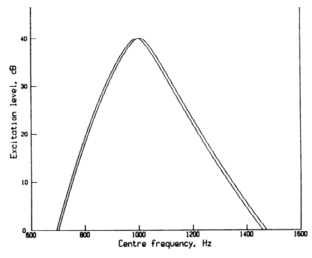


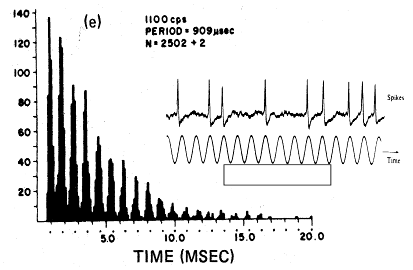
FIG. 5.2 Schematic representation of the patterns of excitation evoked by two tones of slightly different frequency; the frequencies of the tones are 995 and 1005 Hz and their level is 40 dB. The greatest difference in excitation level for the two patterns occurs on the steeply sloping low-frequency side.

A just detectable pitch change at 3 kHz and below leads to a change in excitation level that is too small to be detected.

Therefore - acuity for pitch differences for low frequency sinusoids cannot be explained by place cues.

What other cues are there?

Neural temporal coding



Interval histogram from recordings of auditory nerve responses to 1100 Hz sine wave. The common intervals are at 1/1100 seconds, 2/1100 seconds, etc.

Synchrony of nerve firing times to sine-wave period: very precise up to about 1.5 kHz – then declines and is lost at 5 kHz and above – so timing cues to pitch decline in accuracy above 1.5 kHz

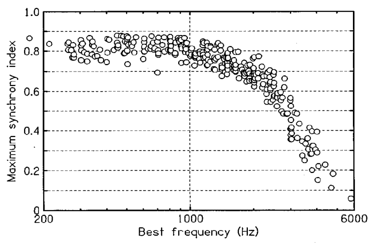


FIG. 10. Synchrony coefficient for 315 neurons as measured by D. H. Johnson who provided data for this plot (courtesy AIP Press).

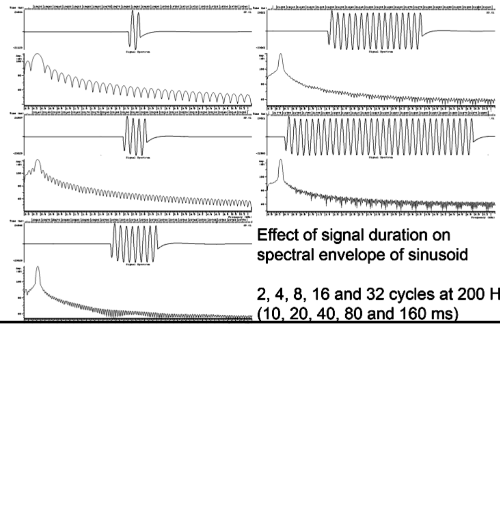
What about effects of duration?

- If pitch discrimination is based on time intervals between nerve firings then as more intervals occur, discrimination is likely to be more accurate in a way that depends on the statistics of timing of nerve firing,

Effects of duration on spectrum

- But duration also affects spectrum, and hence place coding - width of excitation pattern grows with inverse of duration

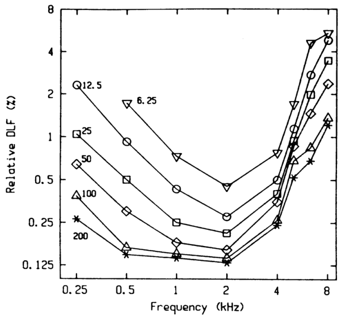
Effects of duration on sine wave spectrum – spectrum spreads at shorter durations which limits place coding of pitch



Effect of signal duration on spectral envelope of sinusoid

2, 4, 8, 16 and 32 cycles at 200 Hz (10, 20, 40, 80 and 160 ms)

Effects of signal duration: place vs. temporal coding of sine wave frequency



Data from Moore

Above 4 kHz there are only place cues – duration has a relatively small effect which can be explained by the spectral spread arising for shorter tones.

Below ~ 4 kHz, pitch discrimination for longer signals is too fine to be explained by place (shifts in excitation pattern)

Effects of signal duration (different curves) are larger at low frequencies. They cannot be explained by spectral of excitation pattern but can be explained by statistics of temporal coding which depends on number of inter-spike intervals.

Coding pure tone frequency

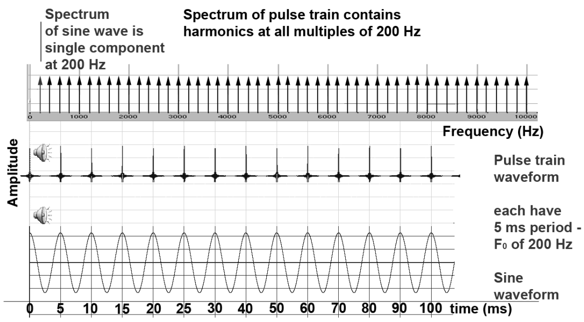
- Only by place of excitation above 4 kHz
- Dominated by temporal coding below ~ 1.5 kHz
- Between 1.5 and 4 kHz both types of cue are available.

20

Pitch of complex sounds

- A complex harmonic sound such as a pulse train has a pitch that is equivalent to that of a sinusoid at the fundamental frequency (F_0) of the pulse signal.
- This information is present in the acoustic signal both in the spectrum, as the frequency of the component at F_0, and in the time domain, as the period of the pulse train.

21



22

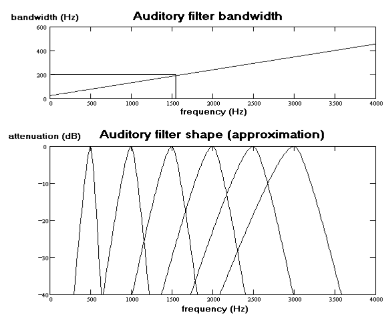
Ohm's other law:

"Every motion of the air, then, which corresponds to a composite mass of musical tones, is, according to Ohm's Law, capable of being analysed into a sum of simple vibrations, and to each such simple vibration corresponds a simple tone, sensible to the ear, and having a pitch determined by the periodic time of the corresponding motion of the air."

(Helmholtz, 1885; "On the Sensations of Tone"

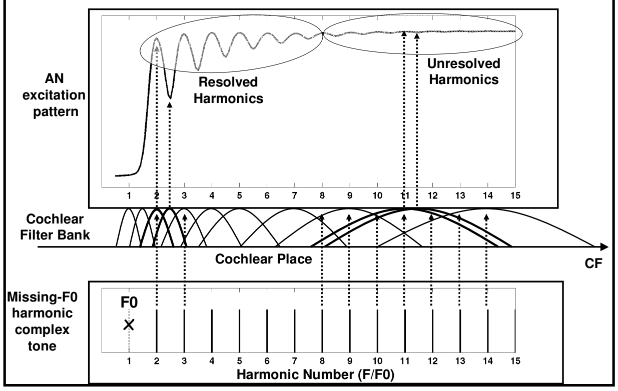
23

Auditory filter bandwidth increases with frequency (while harmonics are evenly spaced). For F_0 of 200 Hz, bandwidth exceeds harmonic spacing above about 1.6 kHz



24

Cochlear frequency selectivity and resolution of harmonics



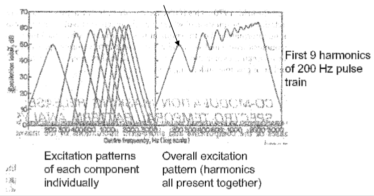
Excitation patterns: complex sounds

Lower harmonics are clearly resolved – For 200 Hz F_0, above 1.6 kHz filter bandwidth is wider than 200 Hz spacing between harmonics and these higher harmonics are not resolved.

Similar limits apply at other F_0s

Classical Place account of pitch

- Pitch of a complex sound determined by position of peak in excitation pattern due to basilar membrane response to fundamental frequency (F_0) component



The missing fundamental

- Schouten (1938, 1940) made a crucial test of the place theory that is based on Ohm's Law
- He presented a pulse signal, with a complete harmonic series. A place account would claim that the pitch is due to the lowest frequency component, at the fundamental frequency.
- This signal is compared to a signal modified to remove the fundamental frequency component. According to place theory, the pitch should change

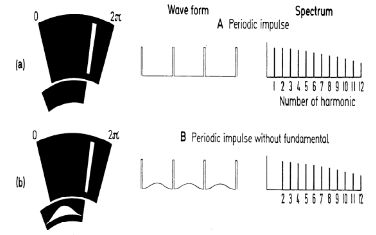


Fig. 7. (a) Periodic pulse series — waveform and spectrum. (b) Periodic pulse series with the fundamental component suppressed. The left-hand part of the figure shows the two masks used in the optical siren. After SCHOUTEN (1940a)

Audio demonstration from "Audio Demonstrations on Compact Disc (ASA 1989).

The first sound is a 200 Hz harmonic complex tone comprising the 1st 10 harmonics. Succeeding sounds have the 1st, 1st and 2nd, 1st thru 3rd, and then 1st thru 4th harmonics deleted.

For most listeners, pitch is unaffected by deletion of harmonic at fundamental frequency

Schouten called this "residue pitch" – attributing the low pitch percept to the periodicity shown in the auditory nerve response to the unresolved higher harmonics

Auditory frequency analysis of a pulse train

Higher harmonics are closely spaced relative to filter bandwidths and are not resolved. The filter output shows the fundamental periodicity of the pulse train

Lower harmonics are completely resolved (1st 5 to 8 harmonics depending on F_0)

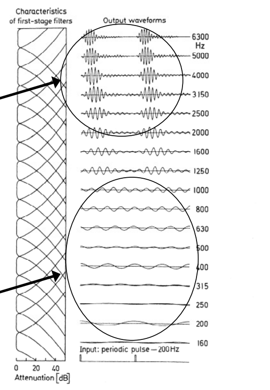


Fig. 11. Periodicity of signals and aural resolution of components: a more modern representation. After PLOMP (1966)

Role of auditory non-linearity?

- Additional frequency components are introduced when a signal is passed through a non-linear system – for harmonic complex tones this could include a distortion component at F_0.
- Can a component introduced at the fundamental frequency explain "The case of the missing fundamental"?

31

Is distortion product responsible for low pitch?

- Patterson (1976) Low frequency noise will mask a distortion component at F_0 – (e.g. a difference tone arising from two adjacent harmonics)
  - but LF noise does not mask the low pitch at F_0
  - therefore the low pitch is not due to distortion

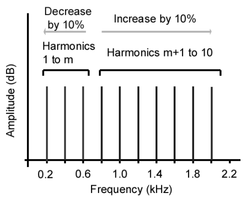
Audio demo – A simple melody is heard played by a series of sine waves and complex tones comprising 3 higher harmonics with the same F_0 as the sine wave. Both the sine and complex tones sound the same melody. Then a low pass noise is added – this masks the sine wave and would mask any auditory distortion product at F_0. The low pitch is still heard from the complex tones.

32

Contributions of resolved and unresolved harmonics

The pitch of the residue suggests that higher **unresolved** harmonics are important in determining the pitch of complex tones. Both Ritsma and Plomp in 1967 published studies that challenged this.

Plomp used stimuli in which the higher and lower harmonics were shifted in frequency in opposite directions. E.g., Harmonics 1 to 4 were shifted down by 10% and harmonics 5 upwards were shifted up by 10%.



33

Contributions of resolved and unresolved harmonics

Generally, and especially in the speech F_0 range, it is harmonics 4 to 8 that dominate pitch

At very high F_0 – above 1.5 kHz, the fundamental frequency component is dominant.

Contributions of unresolved high harmonics never dominate over contributions of resolved harmonics.

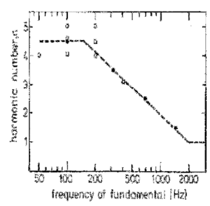


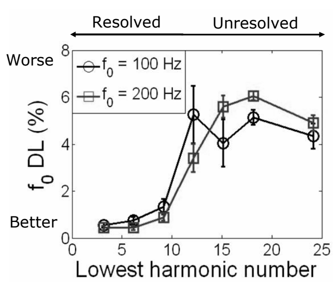
FIG. 46 Harmonics dominant in the sensation of low pitch as a function of the frequency of the fundamental. The solid points are calculated from Fig. 45, the squares from data of Ritsma (1967), the open circles from data of Bilsen (1973).

34

Resolved harmonics produce higher precision of pitch than unresolved harmonics

Harmonic complex tone, 12 successive harmonics

(Bernstein & Oxenham, 2003)



ALSO
Pitch discrimination for complex tones generally better than for the sine wave at F_0 (Henning and Grosberg, 1968)

35

Where are cues to pitch?

The filter output shows the fundamental periodicity – weak cue to pitch

Lower harmonics are completely resolved – their frequencies coded in time (at different places) are primary cues to pitch –DOMINANT AND MOST PRECISE

Harmonic at fundamental frequency not a necessary cue for pitch

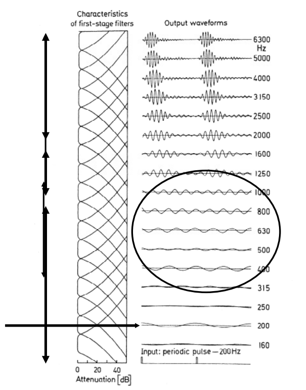


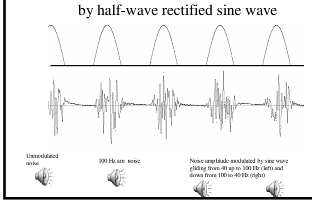
Fig. 11. Periodicity of signals and aural resolution of components: a more modern representation. After PLOMP (1966)

Primary cues for pitch of complex sounds

- Pitch is mostly effectively determined by temporally-encoded representations of the frequencies of resolved harmonics (temporal code needed to explain the precision of pitch discrimination)
- The temporal encoding of F_0 from the unresolved higher harmonics is not a primary cue
- Nor is the harmonic component at F_0 except when F_0 > 1.5 kHz.

Pitch without spectral information

- White noise that is amplitude modulated at rates up to 1000 Hz has a weak pitch (Burns and Viemeister, 1976) The spectrum of the noise is flat, and only temporal cues for the pitch are present
- E.g, below shows white noise (lower trace) amplitude modulated by half-wave rectified sine wave



Purely temporal pitches, although weak, can convey melody information for rates up to 300 or 500 Hz - but very weak above 200 Hz.

Monaural temporal pitch is perceived from the temporal nerve firing pattern, which will be affected by amplitude modulation.

Also DICHOTIC temporal pitches – where a pitch is heard that changes with inter-aural phase.

Current theories of pitch perception

- Pitch perception is based on the pattern of information over a range of frequencies. The major contributing information is the frequencies of the dominant **resolved** harmonics.
- This information is conveyed in the temporal firing pattern of the auditory nerve across frequency channels.
- Pattern processing identifies intervals between nerve firing that are common across frequency channels. For a series of resolved harmonics, nerve firings show a related series of time intervals
- Periodicity information from higher frequency **unresolved** harmonics or from the modulation envelope of noise is another source of input to this pattern processing, but is a relatively weak cue.

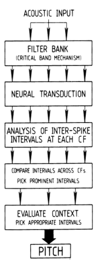
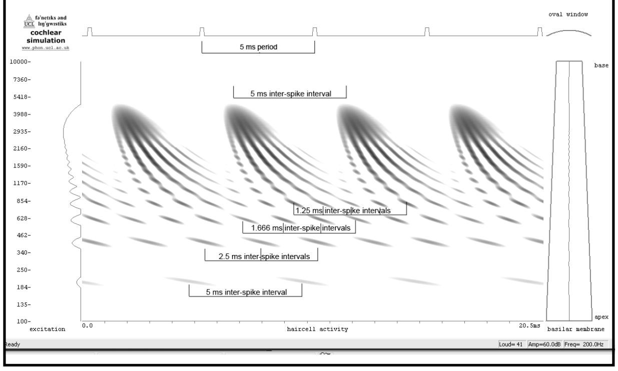


FIG. 5.6 A schematic model for the perception of the pitch of complex tones. The model depends upon both place and timing information.

Figure from Moore and Glasberg (1986)

Auditory nerve responses to pulse train: nerves responding to resolved harmonics show periodicity of each harmonic



Summary: Simple signals

- While pitch is broadly correlated with period, human pitch processing is complex
- Sine waves up to a few kHz - pitch is temporally coded
- Sine waves above 4 kHz, only place cues are present to code sine wave frequency

Summary: Complex signals

- The period indicated by temporal cues alone from unresolved high harmonics in a single auditory filter can signal pitch at $F_{0.}$
  - And a weak pitch can be heard from purely temporal cues with amplitude modulated noise
- However, pitch of complex tones is dominated by resolved harmonics (range 4 to 8 for F_0 in speech range). Here pitch processing depends on pattern extraction operating on time intervals between nerve firings

How might impaired hearing affect pitch perception?

- Wider auditory filters due to OHC damage
  - Fewer harmonics resolved

- Impaired temporal coding
  - Would limit phase-locking and hence temporal coding of frequency
  - Does not seem a major problem in typical SNHL

43