Yet More About Hearing

Josh McDermott
9.35
May 6, 2004

Steven's power law

Based on a number of different experiments, the psychologist S.S. Stevens concluded that loudness could be described as a power law:

$$L = kI^{0.3}$$

(L = Loudness, I = Intensity, k is a constant)

So, a 10-dB increase in level gives a doubling in loudness.

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph. Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.7, p. 341.

Problem: single auditory nerve fibers have narrow dynamic ranges. Only change their firing rate over a 40dB range:

(Image removed due to copyright considerations.)

Solutions to the dynamic range problem

- Off-frequency listening (Siebert, 1970)

(Image removed due to copyright considerations.)

- Temporal information
  - Increase in synchrony with level (limited to frequencies at which phase locking occurs < 4kHz)
- Low- and medium spontaneous-rate fibers
  - Higher thresholds, but greater dynamic range

Intensity discrimination of noise in bandstop noise

(Image removed due to copyright considerations.)

- Viemeister's (1983: Science 221, 1206-8) experiment showed that good intensity discrimination could be obtained even when off-frequency listening and phase locking were unlikely. He explored possible role of low spontaneous-rate fibers.

Accounting for wide dynamic range of auditory system from neural responses

- Single AN fiber has narrow dynamic range
- Use high-threshold, low-spontaneous fibers. *Are there enough and are thresholds high enough*?
- Use fibers with CFs far from tone frequency. *What about broadband noise*?
- Use temporal information (phase locking). *What about high frequencies*?
- Still unclear how intensity is coded…

A bit more on the cochlea - it's not exactly doing Fourier analysis.

(Image removed due to copyright considerations.)

Bandwidths of auditory nerve fibers are much higher for those carrying high frequencies (note log axis). Called "Q filters".

Bandwidth is roughly proportional to frequency.

This means that low frequencies can be estimated more precisely, but that high frequencies can be better localized in time.

Onsets and offsets, even of a pure tone (sine wave) generate lots of other frequencies, especially high ones. So to localize sharp transients (e.g. the beginning of a sound), you need to be sensitive to lots of high frequencies at once.

Another way to think about it:

Just as was the case for natural images, it turns out that natural sounds tend to have 1/f amplitude spectra:

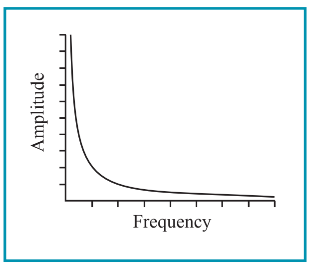

Image by MIT OCW.

So if bandwidth is proportional to frequency, natural sounds will be coded with equal energy in each channel.

Sort of like what we saw in vision.

Harmonicity information is limited by frequency resolution of the cochlea

(Image removed due to copyright considerations.)

1. Resolved harmonics dominate pitch perception (Bird & Darwin, 1998)

(Image removed due to copyright considerations.)

Music Perception

Josh McDermott

9.35

MUSIC

Music is among the most bizarre features of human culture.
-no obvious benefit accrued from making/hearing music
-and yet we love it

Mysteries of music:
Why do people like music?
Why are some things sound musical and others not?
How does music manage to convey emotion?
Why do people like to dance to music?

Are the various features of musical perception innate, or acquired?
If innate, are they an evolved function, or a side effect of something else?

A few basic aspects of music may be hardwired into our brains.

Octave equivalence and pitch perception

Remember the missing fundamental…

If you observe 300, 500, and 600 Hz, there is some ambiguity as to what the fundamental is. It could be 100Hz, but 50Hz would also produce those harmonics.

All scales are based around the octave.

When people with perfect pitch make errors in labeling notes, they generally get the chroma correct but locate the note in the wrong octave.

Because of octave equivalence, the perceptual space for pitch is a helix.

(Image removed due to copyright considerations.)

More evidence for the circularity of pitch comes from Shepard tones:

(Image removed due to copyright considerations.)

The components of a Shepard tone are separated by octaves (e.g. 110, 220, 440, 880 etc.). The frequency of all the components is shifted up or down, but the spectral envelope stays the same.

What is the fundamental?

A sequence of Shepard tones going up the scale sounds like it is forever ascending, even though you just cycle through the same tones over and over.

(Image removed due to copyright considerations.)

The fundamental, and thus the pitch height, is ambiguous.
But you still hear the chroma, and this is what seems to be important in melodies.

Aside on perfect pitch:

It's quite rare (1 in 10,000), but more common in musicians.

Mystery: why doesn't everyone have perfect pitch?

-should be easy to read off pitch from tonotopic representation

Most people seem to represent absolute pitch to some degree.

Absolute pitch may be replaced by relative pitch during development, as a result of exposure to language, music etc.

Perfect pitch probably tells us nothing about music perception, as it is so rare. Still, it is cool.

Consonance and Dissonance

Some notes, when played together, sound good, or natural. We call these consonant.

Others don’t. We call such sounds dissonant.

Pythagoras was the first to note that the notes that sound good together tend to be produced by strings whose lengths are related by simple integer ratios.

Octave: 1:2
Major fifth: 2:3
Tritone: 32:45, called the devil’s interval

What really matters is the fundamental frequency of the note, which is determined by the string length.

Helmholtz deduced that dissonance for complex tones was due to dissonance between their harmonics.

He correctly supposed that dissonance between pure tones was a function of the degree to which they interfere.

Tones close in frequency produce beats - low frequency fluctuations in sound amplitude.

Over time, two pure tones close but not identical in frequency will drift in and out of phase, constructively and destructively interfering. This amplitude modulation is known as beating.

(Image removed due to copyright considerations.)

Consonance increases as two pure tones move apart in frequency.

(Image removed due to copyright considerations.)

The lower tone is 500 Hz.

Thresholds for detecting the tone are measured for different noise bandwidths. The noise is centered on the tone frequency.

(Image removed due to copyright considerations.)

Noise outside a critical band of frequencies does not affect the detection of the tone.

Just like spatial frequencies sufficiently separated from an adapting grating’s freq. were unaffected by the adaptation (remember the bite out of the CSF).

Dissonance is greatest when two frequencies are within a critical bandwidth of each other, i.e. when they produce overlapping waves on the cochlea.

(Image removed due to copyright considerations.)

When two frequencies are more than a critical band apart, they do not interact, at least not in the auditory periphery, and do not produce dissonance.

Complex tones whose fundamental freqs are related by complex ratios have lots of different harmonics that interfere with each other.

Consonant and Dissonant Stimuli

(Image removed due to copyright considerations.)

Pairs of harmonic complex tones forming consonant musical intervals

(Image removed due to copyright considerations.)

The consonance of pairs of complex tones is well predicted by the interference between their harmonics:

(Image removed due to copyright considerations.)

The beating of dissonant stimuli is reflected in IC responses

(Image removed due to copyright considerations.)

Dissonance is reflected in average rates and rate fluctuations of IC neurons

(Image removed due to copyright considerations.)

So there is a physical correlate of perceived dissonance.

This doesn't explain why dissonance sounds bad to most people.

We don't know whether this is hard-wired or acquired.

The Skymaze

(Image removed due to copyright considerations.)

This is a method to measure sound preferences in animals.

A cottontop tamarin

(Image removed due to copyright considerations.)

Using an analogous method, humans show the expected preference for consonance over dissonance:

(Image removed due to copyright considerations.)

Tamarins, however, show no preference at all.

Humans also show a robust preference for white noise over a terrible screeching sound:

(Image removed due to copyright considerations.)

Again, tamarins show no preference. Preferences for sounds may therefore be unique to humans, and perhaps is part of an adaptation for music…