**Pitch Perception**

Oxenham, 2012:

Pitch is one of the primary sensations of sound, among loudness and timbre. It is what defines melody when encountered sequentially and while simultaneous combinations of it define harmony. Pitch is believed to be autocorrelation, meaning the correlate of the periodicity of a waveform. A harmonic complex tone is a periodic waveform whose period corresponds to the fundamental frequency and can be decomposed into sinusoidal harmonics with frequencies at integer multiples of the fundamental frequency. The relative amplitudes of the component harmonic determine another primary sensation, that of timbre, meaning the quality of the sound. Two tones with the same fundamental frequency are perceived as having the same pitch with differences in loudness and timbre.

It is believed that different sound frequencies stimulate different regions of the cochlea. This is called place theory, or frequency to place mapping or tonotopy and carries through to neural representation of sound. Auditory filter bandwidths within the cochlea increase with characteristic frequencies, which means that regions responsible for low-numbered harmonics will respond nearly entirely to one harmonic, while the regions tuned to high-numbered harmonic frequencies will respond to numerous harmonics. The harmonics represented by the former regions are called resolved, while the ones represented by the latter, unresolved.

Yost, 2009:

For a simple, single frequency sound, its frequency is its pitch. The frequency of the spectral component, which is its spectral component at a certain frequency, with certain magnitude and a starting phase, is its perceived pitch. The temporal component of a sound is a sinusoidal time-pressure waveform, for which the reciprocal of the period of the waveform is its pitch. The time-pressure waveform and the spectrum and inverse representations, so the spectrum is the Fourier transform of the time-pressure waveform.

Names some papers that discuss the conflict between spectral and temporal accounts of pitch.

The go on to explain place theory again (if need more content on this check it).

However, past the frequency of 2000 hz, the delineation of the spectral information is not maintained. Since the higher the frequency the lower the resolution of the biomechanical processors (I think they are referring to the hair cells). Low frequencies are represented by single fibres at the apex of the cochlea, while high frequencies are represented by a range of fibres at the base of the cochlea, since these fibres cannot distinguish between small spectral differences. However, it is unclear what constitutes a low and high frequency. Harmonics above the 10th cannot be resolved, an equal difference of 500 hz can be resolved between the 4th 2000 Hz and 5th 2500 hz harmonic (of 500 hz fundamental), while it cannot be resolved between the 10th 5000 hz and 11th 5500 hz harmonic (of 500 hz fundamental). Therefore, it is not a matter of frequency difference rather proportional frequency difference (I made this explanation up).

Relatively fast amplitude changes are the fine structure of the waveform and within the fine structure, slow overall changes in the amplitude are referred to as amplitude modulated envelope of the waveform. A sound pressure waveform can be expressed as the product of the combination of fine structure and envelope. Auditory nerve fibres can respond in a phase locked or synchronous way to fine structure. As the tuned frequency of the fibres increases, the stimulus period decreases and the period of the synchronous firing decreases. If the frequency of the fine structure is higher than 5000 hz, the hair cells cannot respond.

For the sum of multiple harmonics of 200 hz, the time-pressure waveform has an overall amplitude modulated at 200 hz. The amplitude modulation patter is the temporal envelope of the sound. Modulation in neural firing rate can code for a sounds temporal envelope and the reciprocal of the period of the envelope is the perceived pitch.

Models of pitch perception: spectral modelling of pitch perception, temporal modelling of pitch perception (autocorrelation).

Autocorrelation is the correlation of the waveform with itself shifted by a certain time lag t. The autocorrelation model accounts for more pitch data and phenomena compared to other models of pitch perception.

Pitch strength is the perceived difference between the strength of the pitch sensation and that of the timbre sensation.

Oxenham et al., 2011:

Each musical tone is defined by a different pitch, when they are combined they make chords, which are simultaneous pitches and further melodies, which are pitch patterns over time. The pitch of pure tones is their frequency but the pitch of complex tones is the fundamental frequency. The fundamental frequency does not need to be present for the pitch to be perceived (pitch of missing fundamental). At low frequencies, spikes within the auditory nerve occur at multiples of the tones period.

Results from this experiment suggest that the commonly accepted limit of phase locking at 4-5kHz might not be representative of the higher possible perceivable pitch. With higher frequency tones eliciting pitch at the fundamental frequency, strong enough to enable accurate melody discrimination. Therefore, the previously used argument that musical instruments stop at notes around 4-5kHz is not valid as evidence of pitch being perceived up to the limit of 4-5kHz.

The results suggest that pitch can be discriminated by harmonics above 6kHz. Possible methodological explanations for this difference in findings compared to earlier research are posited. For example, previous studies used complex tones of 3 harmonics while this study used at least five. Fewer harmonic complex tones create weaker pitch compared to complex tones made up of a larger number of harmonics. Another explanation could be the use of objective performance measures unlike the previous use of subjective metrics like the participants beliefs regarding if they could discriminate pitch. Possibly the incorporation of noise to the tones could have enhanced the phenomenon of the pitch of the missing fundamental leading to perceptual filling in of the missing fundamental frequency components. Finally, the current study included exclusion criteria for participants whose absolute threshold was above 16Hz. This was not included in the past and therefore limited audibility could have resulted in the different findings.

Wong et al., 2012:

Some languages are tone languages meaning pitch is employed to convey word meaning as well as intonation (speech melody) to convey sentence meaning. Non-tone languages only used pitch through intonation only.

They found that participants form Hong Kong outperformed participants from Canada in a melodic pitch perception task but not a rhythm task. Evidence also suggests that they did not differ in the ability to attend to smaller pitch differences in the context of music.

Evidence points towards speakers of a tone language having enhanced ability in musical pitch processing, when controlling for age, education and musical training. No difference in rhythmic perception. Hong Kong amusics performed better than Canadian amusics at musical pitch perception while controlling for musical training.

The advantage of tone language speakers isn’t due to higher musical advantage (as rhythm wasn’t something they performed better at) or due to higher ability to discriminate between small pitch differences. Rather its because of integrating pitch information across a melody.