Comments

Comment by Carlyon:

One of the attractive aspects of your model is that it can account for the poor rate discrimination by cochlear implantees, despite seemingly accurate phase locking at the level of the auditory nerve (at least when measured in animal experiments). However, it seems less obvious that it can account for the larger effect of duration on frequency discrimination for low-frequency pure tones compared to those of a moderate-to-high frequency (this large duration-dependence is also observed for unresolved harmonics of a low F0). Have you tried to model this aspect of the data?

Reply:

The spectro-temporal model of pitch proposed by Shamma (1985), which we discussed in our oral presentation but not in the written paper, relies on sharp cochlear frequency selectivity giving rise to rapid phase transitions at the cochlear locations of resolved harmonics. Because detection of these phase transitions from neural responses requires phase locking, it seems unlikely that this model could account for the pitch of pure tones at frequencies above 4-5 kHz. Therefore, this model by itself is not likely to account for the different effects of duration on frequency JNDs at low vs. high frequencies. To our knowledge, no pitch model quantitatively accounts for this effect.

Shamma, S.A. (1985). Speech processing in the auditory system I: The representation of speech sounds in the responses of the auditory nerve. J. Acoust. Soc. Am. 78, 1612-1621.

Comment by Yost and Meddis:

We believe that you have been too “harsh” on the use of autocorrelation models to account for the pitch saliency based on resolved vs. unresolved spectra. There have been several successful attempts to account for pitch strength/saliency of complex sounds using autocorrelation (e.g., Yost 1982; Meddis and O’Mard, 1997). It may be true that the height of the major autocorrelation peak of the stimulus cannot account for the weaker pitch saliency of sounds with unresolved components. However, the studies cited above use other aspects of auditory processing of complex sound along with autocorrelation to account for pitch saliency. Thus, we wonder why your paper “stressed” the fact that autocorrelation models cannot account for the weaker pitch saliency of sounds with unresolved harmonics.

Meddis, R. and O'Mard, L. (1997). A unitary theory of pitch perception, J. Acoust. Soc. Am. 102, 1811-1820
Yost, W.A.(1982). The dominance region for the pitch of ripple noise, J. Acoust. Soc. Am. 72, 416-426

Reply:

Our presentation referred to two specific weaknesses of the pooled interspike interval (a.k.a. summary autocorrelation) model of pitch: (1) it underestimates the pitch salience of pure tones compared to complex tones (Cariani and Delgutte 1996); (2) it does not account for the lower salience and poorer discriminability of pitch based on unresolved harmonics compared to that based on resolved harmonics when the harmonics occupy the same frequency region (Shackleton and Carlyon 1994). These conclusions were based on using the largest peaks in the pooled interval distributions (relative to background) as a measure of pitch salience. Yost and Meddis suggest that alternate measures of pitch salience might better account for the psychophysical data.

As pointed out by Carlyon in his comment (below), the autocorrelation in the Yost (1982) model is derived by Fourier transformation of a smeared power spectrum. The mathematical relationship of this "autocorrelation" to the time-domain summary autocorrelation is quite complex, so that conclusions that apply to one representation do not necessarily apply to the other. Even if these two representations can be compared, the rather elaborate pitch salience measure used by Yost (1982) works remarkably well for ripple noise stimuli, but has not been shown to be effective for a wide class of stimuli including pure tones, complex tones with resolved and unresolved harmonics, and AM noise.

Meddis and O'Mard (1997) do not specify a measure of pitch salience *per se*, but rather use the Euclidian distance between the summary autocorrelation functions as a measure of discriminability between two stimuli. The failure of this approach to account for the Shackleton and Carlyon (1994) results has been thoroughly and convincingly documented by Carlyon (1998; see also his comment).

A fundamental issue with the metrics used by Yost (1982) and Meddis and O'Mard (1997) is that, because they depend on the entire shape of the autocorrelation function rather than just the largest peaks at the pitch periods, they are likely to be sensitive to timbre as well as pitch. For this reason, the simpler peak-to-background measure used by Meddis and Hewitt (1991), Cariani and Delgutte (1996), Yost *et al*. (1996), Yost (1996) and in our paper is likely to be more generally useful for pitch salience. Using this measure, we find that pitch salience derived from pooled interspike interval distributions is maximum in the F0 range below 400 Hz where harmonics are not resolved by the cat cochlea, contrary to the Shackleton and Carlyon (1994) results. In view of these results, a search for alternative models of pitch is likely to be productive.

Cariani, P.A. and Delgutte, B. (1996) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J. Neurophysiol. 76, 1698-1716.

Carlyon, R.P. (1998) Comments on "A unitary model of pitch perception" [J. Acoust. Soc. Am. 102, 1811-1820. (1997)]. J. Acoust. Soc. Am. 104, 1118-1121.

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

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

Shackleton, T.M. and Carlyon, R.P. (1994) The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. J. Acoust. Soc. Am. 95, 3529-3540.
Yost, W.A. (1996) Pitch strength of iterated rippled noise. J. Acoust. Soc. Am. 100, 3329–3335.
Yost, W.A., Patterson, R.D., and Sheft, S. (1996) A time domain description for the pitch strength of iterated rippled noise. J. Acoust. Soc. Am. 99, 1066–1078.
Yost, W.A. (1982) The dominance region for the pitch of ripple noise. J. Acoust. Soc. Am. 72, 416-426.

Comment on the comment by Yost and Meddis, by Carlyon

The model by Meddis and O'Mard that is referred to above predicts an effect of frequency region, but not of resolvability (or, as Oxenham's presentation suggests, harmonic number). When frequency region and F0 are varied orthogonally, so that resolvability can be varied within a given frequency region, Meddis and O'Mard's model predicts sensitivity that is not systematically higher for resolved harmonics (Carlyon 1998; Bernstein and Oxenham 2003). It may be that modifications to this model – such as varying the range of lags analysed in each channel to match CF – could eventually account for the data. However, the current version of the model fails on this point, and so it is worth pursuing alternative approaches.

It is also worth noting that the Yost model is very different from the autocorrelation-type models that are currently popular. It starts by calculating the power spectrum, which is then modified by a set of weighting functions derived from nonsimultaneous masking, and these modified spectra are then Fourier transformed. It seems unlikely that this approach could account for the effects of phase on the pitch of unresolved harmonics.

Carlyon, R. P. (1998) Comments on "A unitary model of pitch perception" [J. Acoust. Soc. Am. 102, 1811-1820. (1997)]. J. Acoust. Soc. Am. 104, 1118-1121.
Bernstein, J. and Oxenham, A. (2003) Effects of relative frequency, absolute frequency, and phase on fundamental frequency discrimination : Data and an autocorrelation model. J. Acoust. Soc. Am. 113, 2290.