

HST.723J, Spring 2005 – Theme 4 Report

Madhu Shashanka, shashanka@cns.bu.edu

Introduction

The theme of this report is *pitch and temporal coding*. Pitch, one of the primary auditory percepts, is related to the temporal regularity or periodicity of a sound. It has been the focus of many research studies and its importance stems from its role in perceiving the prosody of speech, melody of music, and in organizing the acoustic environment into different sources. This report is a discussion of five papers which deal with various issues involved in understanding pitch perception. Dai (2000) examines the relative importance of different harmonics in determining the pitch of a harmonic complex tone. Darwin et al. (1995) provide evidence which show that sequential auditory grouping constraints apply to the perception of pitch. Krumbholz et al. (2003) present results of a MEG study which uses a novel approach to identify auditory cortical areas involved in pitch processing. Penagos et al. (2004) examine the representation of pitch salience in auditory cortex using fMRI. Shamma and Klein (2000) propose a model for the formation of harmonic templates from basic properties of early auditory processing.

Psychophysical Studies

Various studies have shown that harmonics in a harmonic complex tone do not make equal contributions to pitch perception. In Dai (2000), the author assumes that pitch is a weighted average of estimated fundamental frequencies based on individual harmonics and sets out to estimate the weighting functions. In the experiment, subjects were asked to judge which of two complex harmonic tones had a higher pitch. The two complex tones on each trial were identical except for random jitters added to individual components (2% jitter at each harmonic). Using a correlation technique, the weight associated with each harmonic was estimated as the correlation coefficient calculated between the frequency perturbations added to that harmonic and the listener's binary responses in the pitch comparison task. From the results, the author concluded that the estimated dominance region in pitch perception had a fixed width in harmonic number (three to four) and a fixed location in absolute frequency (600 Hz) and disagreed with suggestions that the dominance region was roughly constant in harmonic rank. This conclusion would have been more convincing if there had been data from more subjects.

Researchers have studied if cues involved in primitive grouping mechanisms - mistuning, onset time, common modulation, lateral position are a few examples - have any influence on pitch perception. While most such studies dealt with simultaneous perceptual organization, Darwin et al. (1995) tested whether pitch perception was also subject to grouping mechanisms that are concerned with sequential organization. A potent way to perceptually remove a single frequency component from a complex is to embed the complex in a short sequence of tones at the same frequency (sequence condition). The authors wanted to see if a preceding tonal sequence was able to perceptually remove a mistuned harmonic from a complex so that it made less of a contribution to the pitch of that complex. They found that pitch shifts in the sequence condition were smaller than those in the isolated condition (figure 1) and the side of presentation of the mistuned component did not influence the size of pitch shifts. These results show that sequential context does have an influence. An alternative explanation could be that the neural representation of the mistuned component was reduced during the complex by peripheral adaptation. Though many studies have shown this possibility to be unlikely, the authors conducted another experiment to provide further evidence against adaptation effects. They examined the effect of varying the level of a mistuned harmonic on the contribution to pitch and found that reducing the level reduced the size of pitch change only if the component was presented ipsilaterally (figure 2). The reduction in the ipsilateral case is probably due to ipsilateral masking by neighboring components on the mistuned one. The mistuned component will be more prominent with contralateral presentation. This result argues against adaptation effects and against models of pitch perception which weight resolved components independently of the degree to which a harmonic is resolved.

Neuromagnetic/Neuroimaging Studies

Most neuromagnetic studies on pitch processing focus on the prominent N100m response. But it has been shown recently that any components of the N100m associated with pitch are fundamentally confounded with components that reflect other stimulus features such as loudness and timbre. Krumbholz et al. (2003) present a novel approach which enabled them to avoid confounding the pitch-onset response (POR) with the sound-onset response (SOR). They examined the activity evoked when an on-going sound without pitch (random noise) changed to a sound with pitch (regular interval (RI) sound). Both sounds had the same energy and very similar spectral profiles. The prominent perceptual change at the transition was accompanied by a prominent deflection in the magnetic field (POR) (see figure 3, left panel (b)) but there was no discernable response when the transition was from a tone to a noise. According to the authors, this suggests that the cortical generators of POR are associated with neural processing of pitch-related information. They showed that the amplitude and latency of the POR varied with pitch and pitch strength (figure 3, right panels). With further measurements and analysis, the authors suggested that

the location of the POR dipole was on the Heschl's gyrus, anterior and inferior to the N100m dipoles. They also conducted a psychophysical experiment to measure the perceptual integration time for pitch and concluded that it was comparable to the POR latency suggesting that POR corresponds to higher levels of pitch processing.

Previous functional brain imaging work in humans has shown that the level of population neural activity in centers throughout the auditory system is related to the temporal regularity of a sound, suggesting a possible relationship to pitch. In the study by Penagos et al. (2004), fMRI was used to investigate neural representations of pitch salience using harmonic complex tones with varying degrees of pitch salience but identical temporal regularity. Tones consisted of only high-numbered harmonics (weak pitch) or low-numbered harmonics (strong pitch). Tones with a higher F0 which had fully resolved harmonics in both low and high spectral regions (and hence strong pitch percepts) were used as controls (figure 4). They found that activity in anterolateral auditory cortex was specifically sensitive to pitch salience and not to other differences between conditions. Activation produced by the condition with the weak pitch was the lowest (figure 5, left panel). Noise, which had the lowest pitch salience of all, produced the lowest level of activity when comparisons were broken down by spectral region but in contrast to tonal stimuli, spectral region had a significant effect. They found that activation in the inferior colliculus and cochlear nucleus did not differ significantly across stimulus conditions. Based on the results, they concluded that there was a representation for pitch salience in neural activity levels of anterior nonprimary auditory cortex but no identified representation in subcortical structures.

Neural Modeling

Psychoacoustical models of pitch have long postulated the existence of internalized harmonic templates against which incoming resolved spectra can be compared, and pitch determined according to the best matching templates. The main criticism for such spectral pitch theories is the lack of convincing biological evidence for the existence of these templates or for how they might be generated. Shamma and Klein (2000) propose a model which describes how such harmonic templates may emerge as a consequence of basic properties of early auditory processing, and not of exposure to any special sound stimuli such as harmonically rich speech or music. The basic stages of the model are illustrated in the schematic shown in figure 6. An analysis stage consists of a filter bank followed by temporal and spectral sharpening. The second stage is a matrix of coincidence detectors that compute pairwise instantaneous correlation among all filter outputs. Their simulations show that for any broadband stimulus, high coincidences occur between cochlear channels that are exactly harmonic intervals apart and accumulating coincidences over time results in the formation of harmonic templates for all fundamental frequencies in the phase-locking frequency range. Properties like cochlear frequency analysis, rapid traveling wave delays near the wave's resonance, phase-locking and half-wave rectification are necessary for the

model to work. Other details such as enhanced temporal synchrony and spectral sharpness help in improving the templates in a graded fashion. While elements of the earlier stages of the model have clear biological underpinnings, the anatomy and location of the coincidence matrix, and the identity of its immediate input pathway are speculative. The authors offer possible neural realizations but more physiological data is required to prove/disprove their hypotheses. They also propose a model for the pitch of complex tones with unresolved harmonics (residue pitch) but wonder whether it can withstand a critical quantitative scrutiny of its properties. In conclusion, they present a model which shows how harmonic templates can be formed with broadband stimulation, without the need for neural delay lines and other temporal structures.

Discussion and Conclusions

Pitch is an important percept and has been the subject of many studies in auditory research. The study by Dai (2000) throws up interesting questions about the dominance region (harmonics important for determining the pitch). The author's claim that the dominant harmonics are best described as close to a fixed absolute frequency of 600 Hz instead of fixed harmonic ranks (regardless of frequencies) would have been more convincing if there was data from more subjects. The method used (correlation technique) to estimate relative weights of harmonics was interesting and probably better than the one used in a similar study by Moore et al. (1985). Darwin et al. (1995) presented results of two well designed psychophysical experiments which showed that pitch perception is subject to auditory grouping mechanisms that are concerned with sequential organization. These experiments imply that pitch processing involves higher-level brain centers and cannot be the product of peripheral processing alone. Krumbholz et al. (2003) used a nice paradigm to dissociate the response to sound onset from the response to pitch changes in their neuromagnetic study. They were able to localize this pitch onset response to medial Heschl's gyrus. Their data, along with data from other imaging and MEG studies, support the hypothesis that pitch is extracted and refined in centers progressing laterally along Heschl's gyrus and on out into adjacent areas. Penagos et al. (2004) were able to attribute neural activity on anterolateral Heschl's gyrus to perceptual rather than physical stimulus differences. Though they did not identify representation for pitch salience in subcortical structures, they did not rule out its existence on a finer spatial or temporal scale than provided by fMRI. Shamma and Klein (2000) provided further support for spectral pitch theories which have been very successful in explaining and predicting pitches of complex tones. They proposed a biologically plausible model for the formation of harmonic templates which form a critical part of spectral pitch theories. Pitch, which is an important and compelling auditory percept, involves different parts of the auditory system and more research is required to precisely understand the mechanisms.

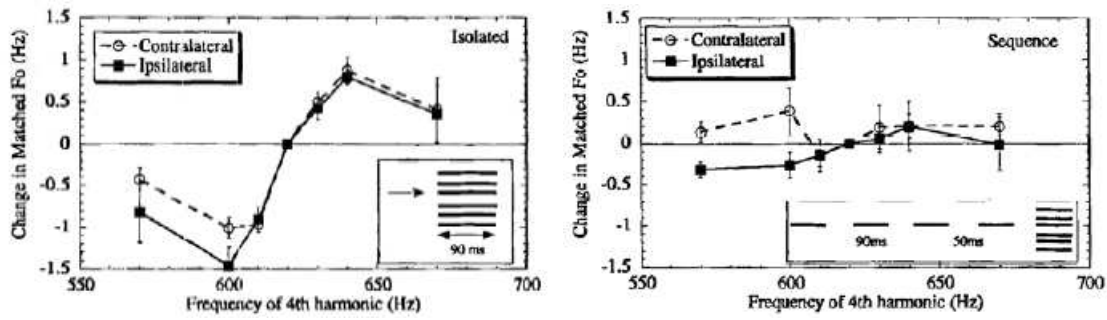


Figure 1: Average pitch matches in isolated (left) and sequence (right) conditions. (From Darwin et al. (1995), figures 2 and 3)

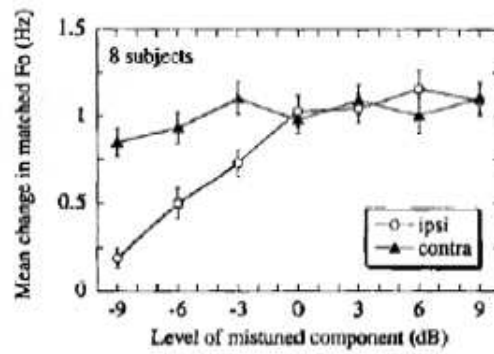


Figure 2: Mean change in matched F0. (From Darwin et al. (1995), figure 7)

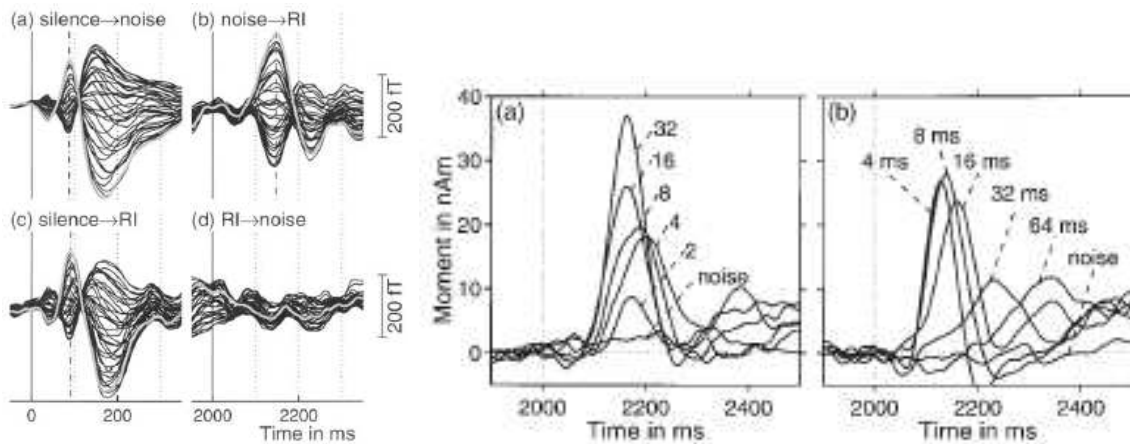


Figure 3: Evoked neuromagnetic fields (left) and average dipole moments as a function of time in response to transition from a noise to a RI sound with fixed delay (a) and fixed no. of iterations (b) (right). (From Krumbholz et al. (2003), figures 3 and 4)

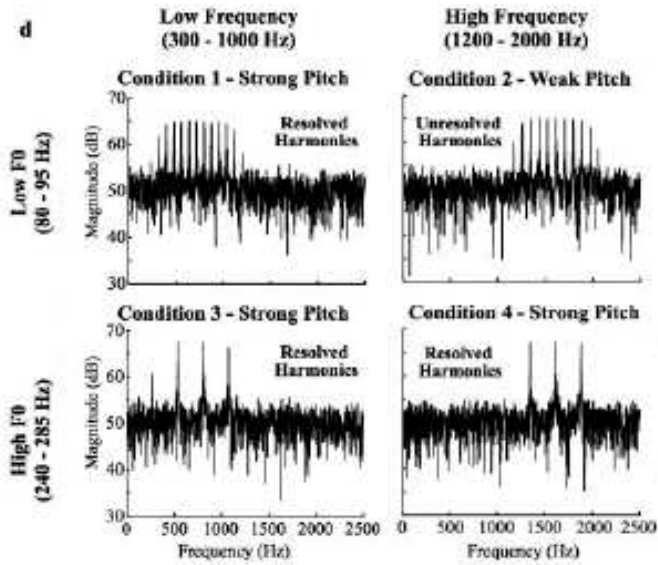


Figure 4: Complex-tone stimuli - different conditions. (From Penagos et al. (2004), figure 1)

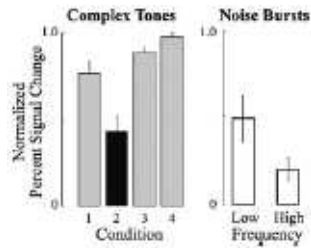


Figure 5: Average percentage signal change for each tone condition (b) or band-pass filtered noise-bursts (c) relative to silence. (From Penagos et al. (2004), figure 2)

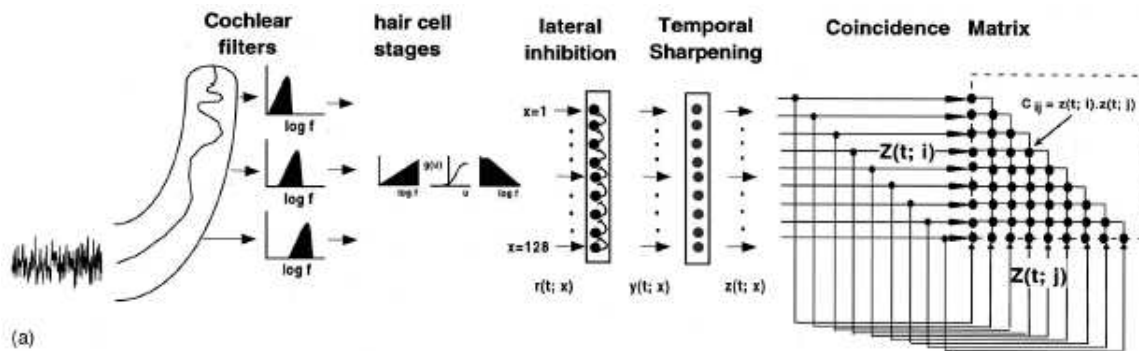


Figure 6: Schematic of the model. (From Shamma and Klein (2000), figure 1)

References

- Dai, H. (2000). On the relative influence of individual harmonics on pitch judgement. *J. Acoust. Soc. Am.*, 107:953–959.
- 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.
- Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., and Lutkenhoner, B. (2003). Neuromagnetic evidence for a pitch processing center in heschl’s gyrus. *Cerebral Cortex*, 13:765–772.
- Penagos, H., Melcher, J. R., and Oxenham, A. J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J. Neurosci*, 24:6810–6815.
- Shamma, S. and Klein, D. (2000). The case of missing pitch templates: How harmonic templates emerge in the early auditory system. *J. Acoust. Soc. Am.*, 107:2631–2644.