Neural correlates of musical dissonance in the inferior colliculus

M. F. McKinney[1,2], M. J. Tramo[2,3] and B. Delgutte[1,2]

[1]*Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary 243 Charles St., Boston, MA, 02114, USA;* [2]*Speech and Hearing Sciences Program, Harvard-MIT Division of Health Sciences and Technology; and* [3]*Harvard Medical School, Dept. of Neurology*
mount@mit.edu

1. Introduction

Complex-tone pairs or chords whose fundamental frequencies are related by "simple" ratios (ratios of small integers) produce a consonant and pleasing sensation while those not related by simple ratios produce a dissonant and rough sensation. Figure 1 shows (A) power spectra of four different musical intervals comprised of pure- and complex-tone pairs and (B) judgements of dissonance for each interval. The interval names are from the Western diatonic scale and, in this case, are all based on A4 (440 Hz). For a given interval, the fundamental frequencies of the complex tones are the same as the frequencies of the corresponding pure-tone interval. For complex-tones, dissonance is highest for the two intervals (minor 2nd and tritone) whose fundamental frequencies are not related by simple ratios (16/15 and 45/32 respectively). For pure-tone intervals, dissonance is also maximum for the minor 2nd but, in contrast to the complex tones, the tritone is not more dissonant than the perfect 4th or 5th.
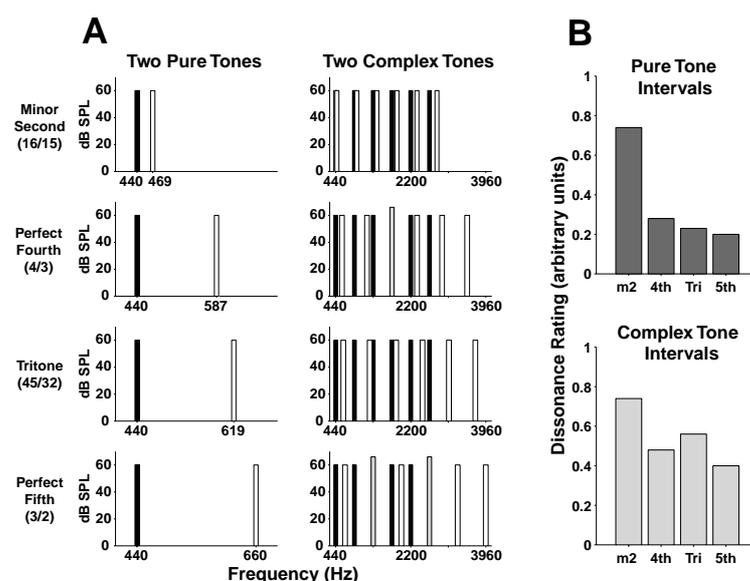

Figure 1. A: Line spectra of pure- and complex-tone intervals based at 440 Hz. The ratios of the fundamental frequencies are given under the interval name. Each tone in a complex-tone pair contains six harmonics, each at the same level as the pure tones. B: Dissonance judgements for two-tone intervals comprised of pure and complex tones (adapted from Terhardt, 1984).

It has been hypothesized that dissonance is caused by beating between neighboring harmonics (Helmholtz, 1863; Plomp and Levelt, 1965). Beating occurs at the frequency difference between two harmonics and can be seen in the stimulus

waveform of the pure-tone minor 2nd in Figure 3A. These amplitude fluctuations (beats) are perceived as *roughness* when they occur at frequencies in the range of 20-200 Hz (Békésy, 1960; Plomp and Steeneken, 1968; Terhardt, 1968a,b; Terhardt, 1974; Vogel, 1974). For complex-tone pairs whose fundamental frequencies are related by simple ratios, less beating occurs overall because some harmonics of the two tones coincide.

Using the stimuli shown in Figure 1A, Tramo et al. (1992; 2000) found a correlate of roughness in the temporal discharge patterns of auditory-nerve (AN) fibers. Their model for roughness operates on AN fibers grouped by characteristic frequency (CF) and uses bandpass filters to extract the temporal fluctuations in each CF band. The filter characteristics were based on the psychoacoustic dependence of roughness on modulation frequency (Terhardt, 1968b). We later noted that these filters resemble modulation transfer functions (MTFs) of inferior colliculus (IC) neurons (Rees and Møller, 1983; Rees and Møller, 1987; Langner and Schreiner, 1988; Rees and Palmer, 1989; Fastl, 1990; Delgutte et al., 1998). Figure 2 shows, for amplitude-modulated (AM) tones, the dependence of roughness on modulation frequency (thick line) along with temporal MTFs from several IC neurons (thin lines). Roughness increases as a function of modulation frequency up to ~70 Hz and then decreases. IC MTFs are slightly more lowpass in shape but, in general, IC neurons strongly respond to stimuli with modulation frequencies in the roughness range. Thus, we hypothesize that there are direct correlates of roughness in the average rates and rate fluctuations of IC neural responses without the need for additional filtering.
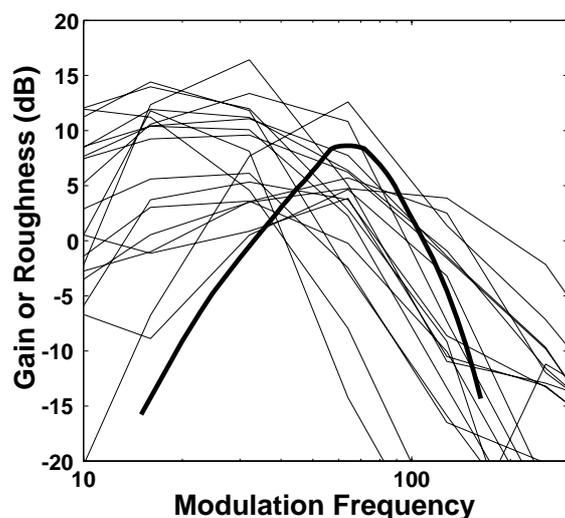

Figure 2. Temporal MTFs of IC neurons (thin lines) and the psychoacoustic rough-ness function (thick line). The MTFs were measured using the procedure of Delgutte (1998) (except with a CF-tone carrier). The psychoacoustic roughness function is for AM tones with a carrier frequency of 1000 Hz (adapted from Fig. 11 in Terhardt, 1968b).

2. Method

Single-unit recordings were made in the IC of Dial-anesthetized cats with tungsten electrodes as described by (Delgutte et al., 1999). Acoustic stimuli were presented diotically unless a stronger response was obtained from monaural presentation. Stimulus levels were typically 60 dB SPL but were adjusted occasionally according to the sensitivity of the neuron. The harmonics for the six-component complex tones were in cosine phase and equal in amplitude. The stimuli included the consonant/dissonant interval stimuli shown in Figure 1A as well as (highly consonant) unison and octave intervals. While the stimuli shown in the figure are based on just intonation, the stimuli used in this study were based on equal temperament tuning.

The average discharge rates and rate fluctuations of the responses were examined for correlates of the relative dissonance ratings shown in Figure 1B. Peri-stimulus time (PST) histograms with 1 msec binwidths were calculated from responses to the musical interval stimuli. A neuron was classified as *Onset* if, for the complex-tone unison interval, the average discharge rate during the first 20 msec of the response was greater than 10 times the rate over the remaining duration of the stimulus. Otherwise the neuron was classified as *Sustained*. Average discharge rates were calculated directly from the histograms and the standard deviation of smoothed (uniformly over 3 bins) histograms was used as the measure of rate fluctuation.
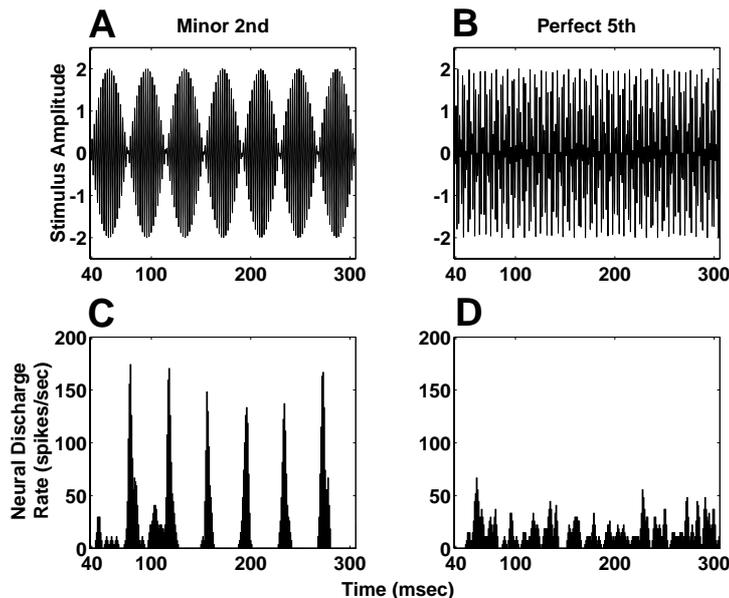

Figure 3. A-B: Minor 2nd and Perfect 5th stimulus waveforms. C-D: Single neuron responses from the IC. CF = 440 Hz.

3. Results

Figure 3C-D shows responses of a sustained IC neuron the pure-tone minor 2nd (A) and perfect 5th (B) musical intervals. The low-frequency beating of the minor 2nd stimulus is clearly reflected in the temporal pattern of the neural response. The interval between firings matches the beat frequency of the stimulus (466 – 440 = 26 Hz). In contrast, the perfect 5th stimulus, which has a much smoother temporal envelope, produces no regular beating in the response.

Figure 4 shows responses of an Onset neuron to all of the pure- (A) and complex-tone (B) musical interval stimuli. The neuron responded only to the onset of pure-tone pairs except for the minor 2nd stimulus. For the minor 2nd it responded over the entire stimulus duration at preferred stimulus phases. The response to complex-tone pairs is slightly stronger than the pure-tone response for all stimuli and the response to the complex-tone minor 2nd also shows a beating pattern. The stronger response to the complex-tone pairs might be attributed to the presence of harmonics near the neuron's CF.

Figure 4C and D show the average discharge rate and rate fluctuations for the Onset neuron in response to all the stimuli. For pure-tones, both response measures show a maximum for the minor 2nd. For complex-tones, both measures also show a peak response at the minor 2nd with a secondary peak at the tritone. These responses are consistent with the dissonance ratings shown in Figure 1B. Thus, for this neuron,

dissonance is reflected in the average rate as well as the rate fluctuations for both pure- and complex-tone pairs.
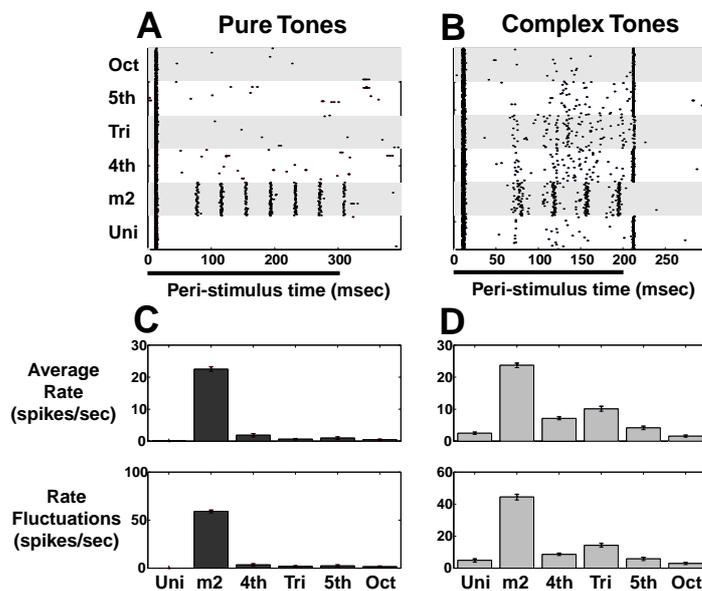

Figure 4. *Response of an Onset neuron to pure- and complex-tone pairs at specified intervals. Upper panels show neural activity over the duration of every stimulus presentation (30 presentations per stimulus). The horizontal black line below each panel indicates stimulus on-time. Lower panels show the average discharge rate and rate fluctuations in the response of this neuron to each interval pair. CF = 1159 Hz. Note the different time scales in A and B.*
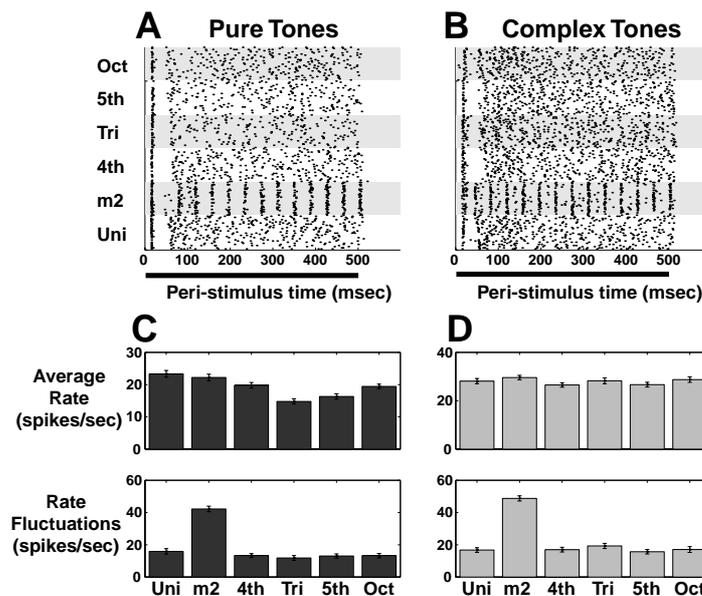

Figure 5. *Same as Figure 4 for a Sustained neuron. CF = 440 Hz.*

Figure 5A and B show the responses of a Sustained neuron to the tone-pair stimuli. In this case, the neuron responds throughout the duration of all stimuli. There is considerable beating in response to the minor 2nd for both pure and complex tones (upper panels) but not so much in response to the complex tritone. Given the CF of this neuron (440 Hz), the response to complex-tone pairs should be dominated by the first pair of harmonics. For the tritone interval, the beat frequency of the first two harmonics is 178 Hz which, based on this neuron's MTF (~ –20 dB gain at 178 Hz), is not likely to be well represented in the response.

The lower panels of Figure 5 show the average discharge rates and rate fluctuations of this Sustained neuron in response to the tone-pair stimuli. The average discharge rates are relatively constant across stimuli and poorly correlate with the psychoacoustic dissonance ratings of Figure 1B. However, the rate fluctuations across

stimulus do correlate with the psychoacoustic data for both pure- and complex-tone pairs (with the exception of the complex-tone tritone). Thus, for this Sustained neuron, rate fluctuations better represent dissonance than do average discharge rates.

Figure 6 shows the mean rate fluctuations across all neurons and the average rate for all Onset neurons in response to the tone-pair stimuli. In all panels, the peak response occurs for the minor 2nd stimulus and, for complex-tones, the tritone elicits the second largest response. These results are qualitatively similar to the psychoacoustic data on dissonance (Figure 1B). An aspect of the neural response that is not seen in the psychoacoustic data is the slight increase in average rate of the Onset response for the pure-tone octave stimulus (panel C). This is due to the fact that most neurons from which we recorded had CFs higher than the pure-tone pair frequencies so they responded more when the upper tone in the pair approached the CF.
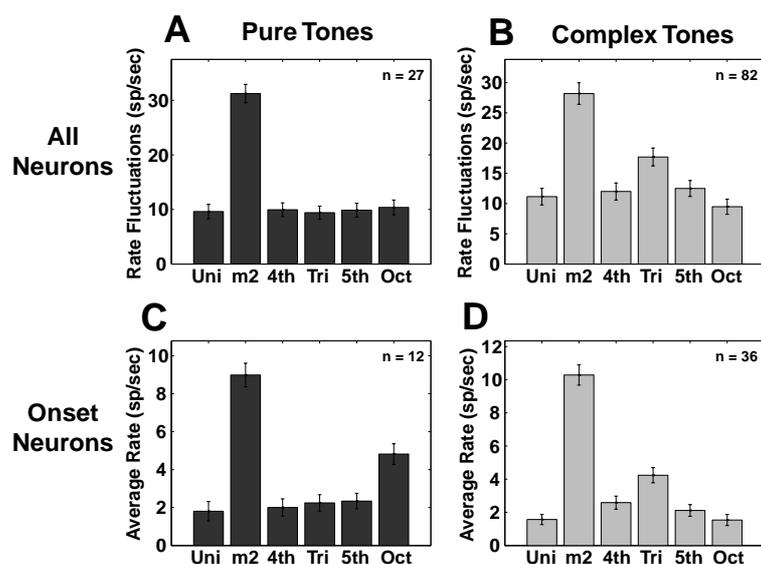

Figure 6. Mean rate-fluctuation across all neurons for the pure- (A) and complex-tone (B) pairs. Average discharge rate across all Onset neurons for the same stimuli (C and D). Error bars are estimated standard errors based on the sum of the variances of the responses from individual neurons.

4. Discussion

Our main finding is that dissonance correlates qualitatively with rate fluctuations of IC neurons as well as with average discharge rates of Onset IC neurons. Although the response of any single neuron may not reflect the dissonance of every stimulus (e.g., the complex-tone tritone in Figure 5B), the relative dissonance of all musical intervals studied here is well represented in the *average response* of IC neurons. Onset neurons in particular are well suited to code for dissonance because dissonance is reflected in both their average rate and rate fluctuations. Overall, rate fluctuation rather than average discharge rate of IC neurons is a more robust coding scheme for dissonance because it correlates with perceptual dissonance ratings for both Onset and Sustained neurons.

In music, the dissonance or consonance of a particular chord has both a contextual component and a sensory component. Terhardt separates *musical consonance* into harmony (contextual) and sensory consonance (1977; 1984). In the context of isolated intervals, harmony is absent and musical consonance and sensory consonance are the same. It is in this context that neural correlates of musical consonance (dissonance) were sought in this study.

The work here is a natural extension of the studies of Tramo et al. (1992; 2000), who found correlates of dissonance in AN discharges. The main difference between

the correlates found in the AN and those described here is that, while additional bandpass filtering is required of the AN responses, direct correlates of dissonance are seen in responses of IC neurons. This implies that bandpass filtering of the envelope occurs between the AN and the IC. MTFs have been measured in CN neurons and, on average, are broader and centered at higher modulation frequencies than those from the IC (Frisina et al., 1990; Rhode and Greenberg, 1994; Delgutte et al., 1998). MTFs have also been measured in the lateral superior olive (LSO) and are also relatively broad and centered higher than IC MTFs (Joris and Yin, 1998). Therefore, because MTFs of most IC inputs are broader than those in the IC, it is likely that the additional filtering occurs within the IC itself. Intracellular recordings from the IC show multiple phases of excitation and inhibition which could implement a bandpass filter (Covey et al., 1996; Kuwada et al., 1997).

Although we find correlates of human perception in neural responses of cats, we are not suggesting that these animals sense dissonance in the same way as humans. It is possible, however, that there are general preadaptations for certain aspects of music processing in the mammalian auditory system.

5. Conclusion

Our results suggest that music perception is constrained by neural processing in the auditory periphery and brainstem and that percepts such as roughness and sensory dissonance may be coded in distinct temporal discharge patterns. Our results are also consistent with the idea that neurons in the IC are specifically important for encoding the temporal envelope of sound.

6. Acknowledgements

7. References

Békésy, G. v. (1960) *Experiments in Hearing*. McGraw-Hill, New York.

Covey, E., Kauer, J. A. and Casseday, J. H. (1996) Whole-cell patch-clamp recording reveals subthreshold sound-evoked postsynaptic currents in the inferior colliculus of awake bats. J. Neurosci. 16, 3009-3018.

Delgutte, B., Hammond, B. M. and Cariani, P. A. (1998). Neural coding of the temporal envelope of speech: Relation to modulation transfer functions. In: A. R. Palmer, A. Q. Summerfield and R. Meddis (Eds.), *Psychophysical and Physiological Advances in Hearing*. Whurr, London, pp. 595-603.

Delgutte, B., Joris, P. X., Litovsky, R. Y. and Yin, T. C. T. (1999) Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus. J. Neurophys. 81, 2833-51.

Fastl, H. (1990) The hearing sensation roughness and neuronal responses to AM-tones. Hear. Res. 46, 293-296.

Frisina, R. D., Smith, R. L. and Chamberlain, S. C. (1990) Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement. Hear. Res. 44, 99-122.

Helmholtz, H. v. (1863) *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. F. Vieweg und Sohn, Braunschweig.

Joris, P. X. and Yin, T. C. T. (1998) Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. J. Neurophysiol. 79, 253-269.

Kuwada, S., Batra, R., Yin, T. C., Oliver, D. L., Haberly, L. B. and Stanford, T. R. (1997) Intracellular recordings in response to monaural and binaural stimulation of neurons in the inferior colliculus of the cat. J. Neurosci. 17, 7565-81.

Langner, G. and Schreiner, C. E. (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J. Neurophys. 60, 1799-1822.

Plomp, R. and Levelt, W. J. M. (1965) Tonal consonance and critical bandwidth. J. Acoust. Soc. Am. 38, 548-560.

Plomp, R. and Steeneken, H. J. M. (1968) Interference between two simple tones. J. Acoust. Soc. Amer. 43, 883.

Rees, A. and Møller, A. R. (1983) Responses of neurons in the inferior colliculus of the rat to AM and FM tones. Hear. Res. 10, 301-330.

Rees, A. and Møller, A. R. (1987) Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. Hear. Res. 27, 129-143.

Rees, A. and Palmer, A. R. (1989) Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. J. Acoust. Soc. Am. 85, 1978-1994.

Rhode, W. S. and Greenberg, S. (1994) Encoding of amplitude modulation in the cochlear nucleus of the cat. J. Neurophys. 71, 1797-1825.

Terhardt, E. (1968a) Über die durch amplitudenmodulierte Sinustöne hervorgerufene Hörempfindung. Acustica 20, 210-214.

Terhardt, E. (1968b) Über akustische Rauhigkeit und Schwankungsstärke. Acustica 20, 215-224.

Terhardt, E. (1974) On the perception of periodic sound fluctuations (roughness). Acustica 30, 201-213.

Terhardt, E. (1977). The two-component theory of musical consonance. In: E. F. Evans and J. P. Wilson (Eds.), *Psychophysics and Physiology of Hearing*. Academic Press, London, pp. 381-390.

Terhardt, E. (1984) The concept of musical consonance: A link between music and psychoacoustics. Music Perception 1, 276-295.

Tramo, M. J., Cariani, P. A. and Delgutte, B. (1992) Representation of tonal consonance in the temporal firing patterns of auditory-nerve fibers: Responses to musical intervals composed of pure tones vs. harmonic complex tones. Soc. Neurosci Abstr.

Tramo, M. J., Cariani, P. A., McKinney, M. F. and Delgutte, B. (2000) Neural coding of tonal consonance and dissonance. Abstracts of the 2000 Midwinter Meeting of the Association for Research in Otolaryngology.

Vogel, A. (1974). Roughness and its relation to the time-pattern of psychoacoustical excitation. In: E. Zwicker and E. Terhardt (Eds.), *Facts and Models in Hearing*. Springer-Verlag, New York, pp. 241-250.