

Evaluating Tenney's critical band using a computational model of the human cochlea

Ashkan Fakhrtabatabaie^{1,2†}

Skyler G. Jennings²

¹ College of Fine Arts, University of Utah, Salt Lake City, Utah, USA

² College of Health, University of Utah, Salt Lake City, Utah, USA

† Corresponding author: ashkan.fakhrtabatabaie@utah.edu

Published 16 December 2021; <https://doi.org/10.18061/FDMC.2021.0048>

Author video presentation and/or other conference material: <https://doi.org/10.17605/OSF.IO/TMDFY>

Abstract

To understand and formalize the perceptual outcomes produced by certain orchestration techniques, researchers studying the psychology of timbre rely largely on combining the listeners perception reports with audio analysis of musical excerpts. Despite their ability to predict many aspects of auditory perception, mathematical auditory models have been scarcely explored in understanding the neurophysiological basis of music perception. If implemented, these models can quickly simulate physiological responses thought to mediate the perception of various timbral effects. Using a computational auditory model (Zilany et al., 2014), we present a physiological analysis of an excerpt from *Critical Band* (Tenney, 1988/2000), which uses loudness fluctuations as a structural element to create implicit timbral phenomena such as beat frequency and roughness. In this study, the results of the score analysis, audio analysis, and auditory model predictions were compared. The score analysis (Fakhrtabatabaie, 2020) suggested an implicit timbral component that the audio analysis of a computer realization of the score was not able to demonstrate. However, the auditory model prediction supported the timbral effect observed in the score analysis. The results of this study suggest that the auditory model can successfully provide a neurophysiological correlate of an aspect of perception of a complex musical stimulus. The use of an auditory model in future research may help with interpreting the score and predicting what perceptual experiments are likely to reveal.

KEYWORDS: *critical band, beat frequency, roughness, auditory filter, temporal envelope, inner hair cells, model simulation*

Introduction

Critical Band is a music composition for any 16 or more sustaining instruments (Tenney 1988/2000). The title of *Critical Band* refers to a psychoacoustic concept with the same name, introduced by physicist Harvey Fletcher (1940). Fletcher used auditory masking techniques and established the concept of the critical band (CB) and

auditory filters (AF) to account for the limitation of the auditory system in resolving simultaneous tones with a narrow frequency separation (e.g., 1-5%). Our ears are sensitive to a wide range of frequencies and the basilar membrane (BM), located within the cochlea in the inner ear, plays a crucial role in separating input frequencies. Due to longitudinal variation of mass and stiffness along the BM, spatially separate locations on the BM vibrate independently from each other, and each location responds to a specific characteristic frequency (CF). Thus, the BM functions as a bank of bandpass AFs enabling the resolution of frequency components within complex input signals. The CB quantifies this frequency-tuning property of the inner ear: the CB is the width of frequencies that optimally pass through an AF without being attenuated. However, if two frequencies within an input signal are too close (e.g., 5%), they cannot excite two fully independent CF locations on the BM. In other words, the excitation patterns elicited by these frequencies overlap (i.e., they pass through the same AF). One perceptual result of this overlap is that the excitation of the two frequencies will beat at a frequency equal to the frequency difference of the two components and give rise to the sensation of fluctuations in loudness or a roughness in timbre (Viemeister, 1979).

In the score analysis of *Critical Band*, Fakhrtabatabaie (2020) used intervallic analysis combined with phenomenology approaches and provided a detailed prediction of the anticipated timbral effects upon an accurate performance of the piece. According to Tenney's (1988/2000) program notes, the sensation of beat or roughness generated by simultaneous tones that are within the CB are two of the main timbral effects that Tenney desires to achieve in *Critical Band*. The timbral effects generated by this process are the perceptual results of destructive and constructive interferences of the amplitude envelopes of simultaneous tones within the CB also known as temporal envelope. The difference between the frequencies of the tones defines the frequency of the



beat they generate. If the difference between the frequencies of two tones within the CB increases, the resulting beat become faster. When the frequency of the beat reaches above ~20 Hz the beat is perceived as roughness (Plomp & Steenken, 1968; Terhardt, 1974, 1978). This roughness has been described as the basis for the perception of sensory dissonance (Terhardt, 1976, 1984; Tramo et al., 2001). The roughness disappears and the sensory dissonance resolves as soon as the frequency difference exceeds the CB. In this study we implemented an auditory model to determine if physiological responses from the inner ear were consistent with the anticipated timbral effects in *Critical Band*. The perception of beats and roughness in response to tones within the CB may be associated with spectral distortion produced by inner hair cells (IHC) transduction, which is the transformation of the mechanical energy of the BM to electrical potentials that drive the response of the auditory nerve (Lins & Picton, 1995).

Zilany et al. (2014) introduced a computational model of the human auditory system based on inner ear physiology measured from laboratory animals (e.g., Carney et al., 1993) and psychophysical findings in humans (Moore and Glasberg, 1993; Shera et al., 2002). The model includes stages that simulate middle ear, BM, IHC, synaptic, and auditory nerve function. Model simulations may be realized in MATLAB by inputting any acoustic waveform and obtaining the predicted response for an array of CFs for one or more model outputs including IHC, synapse, and auditory nerve stages. The simulated physiological responses from the model have been rigorously validated against a wide range of empirical data from animal physiology experiments (Zhang et al., 2001; Bruce et al., 2003; Jackson and Carney, 2005; Zilany et al., 2009), and experiments in simple psychophysics (e.g., Jennings et al., 2011) and speech perception (e.g., Wirtzfeld et al., 2017).

We hypothesized that the timbral effects predicted in the score analysis will be revealed in the time-varying response of simulated IHCs from the auditory model.

Method

Using the program notes written by the composer and the interval analysis using frequency figures we (1) found a segment of *Critical Band* in which both beats and timbral roughness emerge, and (2) hypothesized a numerical measurement of beats and timbral roughness

(discussed later and presented in Table I) to test by the physiological analysis.

Because the timbral effects explored in this study are the distortion products created by the nonlinearity of the inner ear we hypothesized that they would not appear in a spectral analysis of the acoustic signal. However, a spectral analysis of IHC responses to the acoustic signal should carry these timbral effects. To test this, a Fast Fourier Transformation (FFT) was employed to analyze the frequency components of both acoustic and physiological signals of Segment 5. A spectral analysis was used – as opposed to other analyses – as such analysis is physiologically realistic, given the known tonotopic representation of sound throughout the auditory system (Pickles, 2013). We performed the analysis on a 4000 ms slice of a digitally realized version of Segment 5 that we simulated using sine waves. This sine wave realization was done to eliminate any performance-related inconsistency in order to render the perceptual result in a condition that the tones were performed as accurately as possible. This slice was used as the acoustic input signal for acoustic and auditory model analysis. The output of an ensemble of 26 IHCs over an array of CFs spanning the frequency range between 125 Hz to 3 kHz responding to this acoustic signal was calculated and used for the physiological analysis.

Results

Score Analysis

Segment 5 of *Critical Band* consists of a five-note chord built out of stacks of 13.8 Hz intervals. The chord tones are G \sharp (412.5 Hz), G \sharp (426.2 Hz), A (412.5 Hz), B \flat (453.8 Hz), and B \flat (467.5 Hz). The beat frequency column in Table 1 shows all possible intervals existing between the tones in Segment 5. Because of the symmetrical construction of this chord the 13.8 Hz interval and its first three harmonics were found four, three, two, and one time respectively. The occurrences column in Table 1 shows the number of times each interval was found in the intervallic analysis. We speculated that a larger number of occurrences should result in a larger amplitude of the physiological response of that beat frequency.

Table 1: All possible intervals (left column) existing between the tones and number of occurrences (right column) in Segment 5 of Critical Band. From Fakhrtabatabaie (2020).

Segment 5	
Beat frequency (Hz)	Occurrences
13.8	4
27.5	3
41.4	2
55	1

Physiological Analysis

Figure 1 shows the simulated waveform of 4000 ms of the audio signal of Segment 5. As you can see, the temporal envelope created by beats of the available tones in this segment is revealed by slower periodic dips and peaks. Figure 2 shows the frequency spectrum of Segment 5 using FFT. As expected, the temporal envelope associated with the implied beat frequencies discussed in the score analysis (13.8, 27.5, 41.4, and 55 Hz) are not present in the frequency spectrum of the acoustic signal. FFT of the acoustic signal can only demonstrate the frequency components that exist in the acoustic signal, but is not capable of showing how the acoustic spectrum is modified by non-linear processes inherent in the peripheral auditory system and how these processes are expected to affect perception.

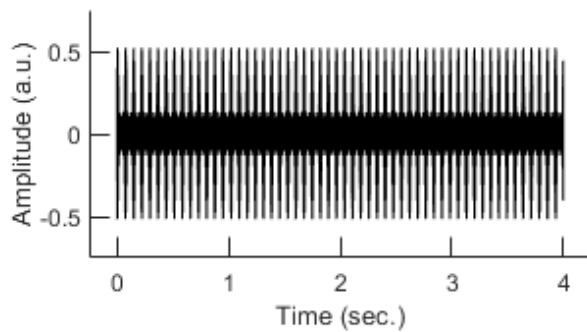


Figure 1: Simulated waveform of 4000 ms of the audio signal of Segment 5.

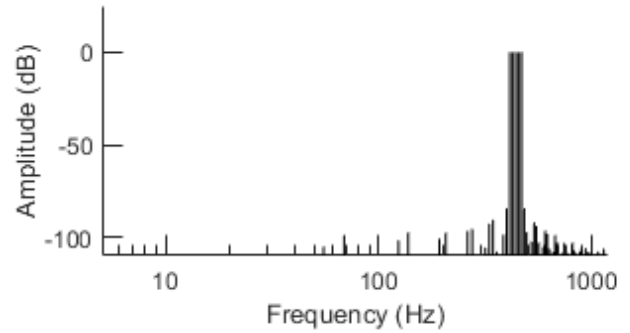


Figure 2: Frequency spectrum of the simulated waveform of Segment 5 calculated using FFT.

Figure 3 shows the output of IHCs of 13 different CFs ranging from 200 to 1000 Hz in response to Segment 5 of *Critical Band*. The output of the IHCs depends on the CF of the patch of BM to which the IHC resides within the inner ear. The IHCs located in close proximity to regions of the BM tuned to the input frequencies generate the strongest output. Thus, Figure 3 shows how only the IHCs located around the area of the BM responding to the characteristic frequencies ranging from 412.5 Hz to 467.5 Hz respond strongly to the inputs from Segment 5. Figure 4 shows the sum of the output signal of all individual IHCs simulated. The output of this population of IHC is very similar to the inputted acoustic signal, with one difference: the negative values of the amplitude fluctuation are clipped (i.e., rectified). This change in the physiological signal results in an output that contains frequency components not present in the acoustic input signal. Figure 5 shows the results of the spectral analysis of the physiological signal corresponding to Segment 5 using FFT, which was identical to the approach we employed in the acoustic signal analysis. The physiological signal contains the predicted beat frequencies from the score analysis. The results are also consistent with our second hypothesis that the number of iterations of each beat frequency is proportional to the relative amplitude of the physiological response. The results are consistent with the notion that nonlinear neural distortion produced by the transduction process of the IHCs contributes to the distinct beat percept experienced when listening to *Critical Band*.

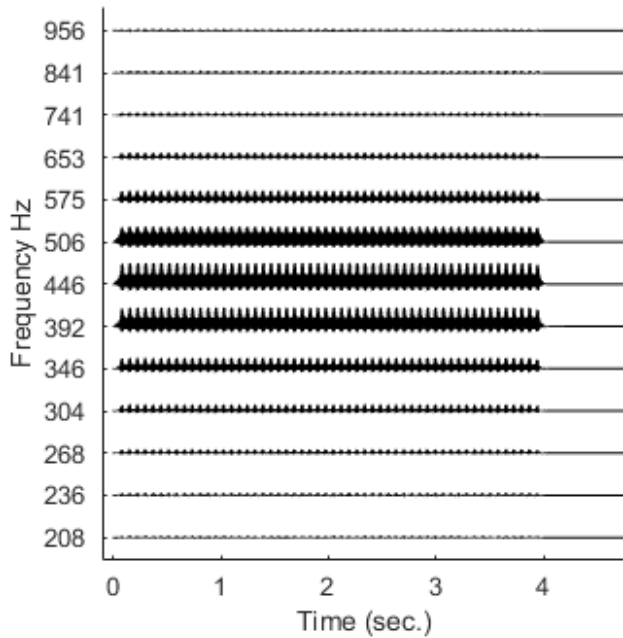


Figure 3: The output of individual IHCs of 13 different CFs ranging from 200 to 1000 Hz responding to Segment 5.

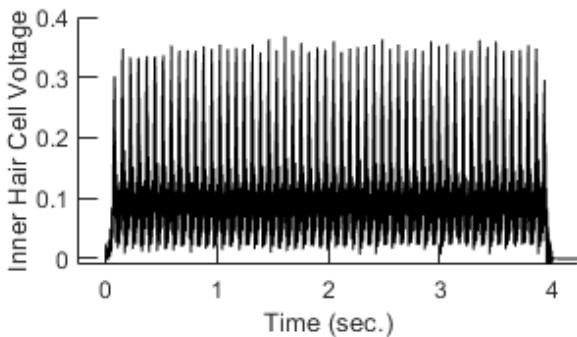


Figure 4: The sum signal of IHCs responding to Segment 5.

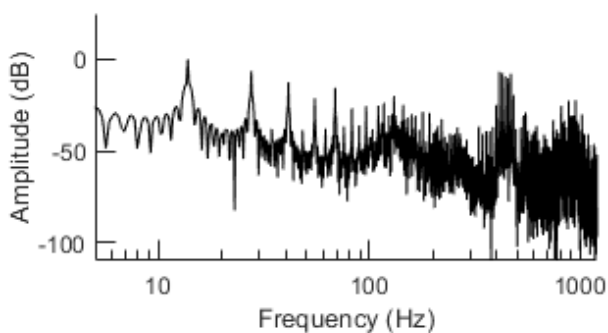


Figure 5: Frequency spectrum of the sum signal of IHCs calculated using FFT.

Discussion

The score analysis of *Critical Band* suggested implied envelope frequency components—resulting in beating and roughness in timbre—that a spectral analysis of the computer realization of the tones in the score was not able to detect. However, the auditory model prediction supported the implied frequency components observed in the score analysis. These results suggest that, if the computational auditory model employed in this study is accurate, the implied beats and timbral roughness suggested in the score analysis are generated in the inner ear and will be carried over to further stages of auditory processing.

Conclusion

A perceptual analysis can help theorists to understand and formalize the sonic outcomes that any acoustic stimulus will generate in the process of sound perception. In listeners’ perception reports, audio segments are played for the listeners directly. Through a perceptual task or a survey that targets the expected perceptual outcomes of the audio segments, the listener’s perception of these outcomes would be analyzed. The listener perception report method has some limitations, such as the lengthy process of data acquisition, the availability of a very controlled testing environment, as well as controlling for the many possible ways that the design of the survey or the perceptual task could create biases that may result in a false perceptual outcome. Our results show that a new model simulation combined with a thorough physiological understanding of a musical phenomenon could be used to predict some aspects of what perceptual studies are likely to reveal.

References

- Bruce, V., Green, P. R., & Georgeson, M. (2003). *Visual perception : Physiology, psychology and ecology*. (4th ed.) Psychology Press.
- Fakhratabaie, A (2020). *Critical Band: A Continuum of Rhythm and Pitch Perception* (Order No. 28153338) [Doctoral disertation, University of Utah]. ProQuest Dissertations and Theses Global.
- Fletcher, H. (1940). Auditory Patterns. *Reviews of Modern Physics*, 12(1), 47–65. <https://doi.org/10.1103/RevModPhys.12.47>
- Franklin, J. (2006). *Settling Scores: A Life in the Margins of American Music*. Sunstone Press.
- Jackson, B. S., & Carney, L. H. (2005). The spontaneous-rate histogram of the auditory nerve can be explained by only two or three spontaneous rates and long-range

- dependence. *Journal of the Association for Research in Otolaryngology*: JARO, 6(2), 148–159. <https://doi.org/10.1007/s10162-005-5045-6>
- Jennings, S. G., Heinz, M. G., & Strickland, E. A. (2011). Evaluating adaptation and olivocochlear efferent feedback as potential explanations of psychophysical overshoot. *Journal of the Association for Research in Otolaryngology*: JARO, 12(3), 345–360. <https://doi.org/10.1007/s10162-011-0256-5>
- Lins, O. G., & Picton, T. W. (1995). Auditory steady-state responses to multiple simultaneous stimuli. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 96(5), 420–432. [https://doi.org/10.1016/0168-5597\(95\)00048-W](https://doi.org/10.1016/0168-5597(95)00048-W)
- Moore, B. C. J. (2012). *An Introduction to the Psychology of Hearing*. Netherlands: Emerald.
- Plomp, R., & Steeneken, H. J. M. (1968). Interference between Two Simple Tones. *The Journal of the Acoustical Society of America*, 43(4), 883–884. <https://doi.org/10.1121/1.1910916>
- Tenney, J. (2000). *Critical Band: 1988/2000*. Lebanon, NH: Frog Peak Music, 2000s.
- Terhardt, E. (1974). On the Perception of Periodic Sound fluctuations (Roughness). *Acta Acustica United with Acustica*, 30(4), 201–213.
- Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds. *Perception & Psychophysics*, 23(6), 483–492. <https://doi.org/10.3758/BF03199523>
- Terhardt, E. (1984). The Concept of Musical Consonance: A Link between Music and Psychoacoustics. *Music Perception: An Interdisciplinary Journal*, 1(3), 276–295. <https://doi.org/10.2307/40285261>
- Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. *Annals of the New York Academy of Sciences*, 930, 92–116. <https://doi.org/10.1111/j.1749-6632.2001.tb05727.x>
- Wirtzfeld, M. R., Pourmand, N., Parsa, V., & Bruce, I. C. (2017). Predicting the quality of enhanced wideband speech with a cochlear model. *The Journal of the Acoustical Society of America*, 142(3), EL319. <https://doi.org/10.1121/1.5003785>
- Zilany, M. S., Bruce, I. C., Nelson, P. C., & Carney, L. H. (2009). A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. *The Journal of the Acoustical Society of America*, 126(5), 2390–2412. <https://doi.org/10.1121/1.3238250>
- Zilany, M. S. A., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *The Journal of the Acoustical Society of America*, 135(1), 283–286. <https://doi.org/10.1121/1.4837815>