

BME 0511 Final Project Proposal | Andrew Sivaprakasam

1. Scientific Question/Biomedical Application:

Despite extensive investigation of both perception and coding of Temporal Fine Structure (TFS) and Envelope (ENV) and how they play a role in speech intelligibility and discrimination, *music* still remains somewhat an enigma. Though multiple researchers have attempted to study music intelligibility and appreciation by creating music-music and music-noise chimaeras, current work consists of studies that are primarily behavioral. Although these studies examine the roles of TFS and ENV in music encoding, their implications may be influenced by higher-level processing beyond the auditory periphery or by auditory filtering effects present in the cochlea.

It is important to study how music is encoded independent of central processing and the relevance of auditory filtering effects in order to drive better design of hearing-assistive devices, such as cochlear implants. It has been established that cochlear implant users, despite developing substantial conversational proficiency, struggle to appreciate music through their implants [1]. To objectively study the encoding and filtering done by the auditory periphery, auditory nerve studies may be utilized.

Advances in modeling and characterization of TFS and ENV using spike-train analyses methods such as the 2018 update to the Zilany et al. model of the auditory periphery paired with alternating-polarity Peristimulus Time Histograms (apPSTHs) allow for better investigation of music coding through simulation. **I propose a model-based investigation of TFS and ENV contributions to coding of a specific musical attribute, *timbre*, and how these contributions are affected by hearing impairment.**

2. Data Set:

The stimuli that I plan to use will be derived from **sound samples from the Philharmonia Orchestra** [2]. This sound sample library has thousands of neatly categorized and concise recordings that span the instrumental range of the standard orchestra, over several pitches, articulations, and expressions. Music-noise chimaeras will be created from a select subset of these recordings, using chimaerizers similar to those used in a previous study [3].

3. Signal Processing Algorithms:

From prior spectral and envelope analysis of timbral characteristics, it is clear timbre is multifactorial and has dependence on both TFS/ENV characteristics of the sound [4]. The proposed project intends to further expand our understanding of this relationship from a neural coding standpoint.

This project requires a good understanding of signal processing ideas and tools covered in this course including the Hilbert analytic signal and spectral analysis. Furthermore, comprehending the advantages outlined by Parida et al. requires learning the limitations of correlogram-based metrics used by Heinz et al. when investigating speech-noise chimaeras using a prior version of the same Zilany model [5]. I will attempt to integrate prior (amateur) knowledge of music and audio engineering with computational modeling and spike-train analysis.

4. Relevant Literature:

My project depends heavily on work by Smith, Bruce, and Parida et al. [3, 6, 5]. The chimaerizers will come from Smith, the auditory nerve model from Bruce, and the apPSTH TFS/ENV methods from Parida. Additionally, I have drawn and will draw inspiration from other papers that have investigated coding of musical ideas and tonal languages from perceptual and neurophysiological perspectives [7, 8, 9, 10, 11, 12].

5. Feasibility Issues:

The success of this project highly depends on smooth integration of the aforementioned work. I will ensure this integration is feasible early on (within the first week of the project). Additionally, even if the same tone (for example A4, $F_0 = 440$ Hz) is studied across multiple instruments, there may be artifacts and onset/offset characteristics that affect the stimulus response that will need to be investigated and either acknowledged or corrected for.

6. Grading:

B Grade: I can demonstrate successful integration of the Philharmonia sound samples/created chimaeras with the BEZ2018 model and the apPSTH analysis, exemplified by a meaningful conclusion on how TFS/ENV contributions to timbral coding are/are not affected by hearing loss in relatively pure tone sounds (i.e. no articulation like dynamic changes/pizzicato/accents) over a few different instruments.

A Grade: I can expand the above to incorporate TFS/ENV contributions to timbral coding in sounds with articulation. I hypothesize coding of articulation characteristics will be highly dependent on ENV contributions while variations in timbral characteristics from instrument to instrument will be dependent mostly on TFS contributions.

7. Other relevant issues:

After outlining this project, it feels ambitious (maybe it isn't?), but interesting. So feel free to suggest ways to scale back if necessary.

References

- [1] Joseph Heng, Gabriela Cantarero, Mounya Elhilali, and Charles J. Limb. Impaired perception of temporal fine structure and musical timbre in cochlear implant users. *Hearing Research*, 280(1):192–200, October 2011.
- [2] Sound samples: <https://philharmonia.co.uk/resources/sound-samples/>.
- [3] Zachary M. Smith, Bertrand Delgutte, and Andrew J. Oxenham. Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876):87–90, March 2002.
- [4] Harin Lee and Daniel Müllensiefen. The Timbre Perception Test (TPT): A new interactive musical assessment tool to measure timbre perception ability. *Atten Percept Psychophys*, 82(7):3658–3675, 2020.
- [5] Satyabrata Parida, Hari Bharadwaj, and Michael G. Heinz. Spectrally specific temporal analyses of spike-train responses to complex sounds: A unifying framework. *bioRxiv*, page 2020.07.17.208330, July 2020. Publisher: Cold Spring Harbor Laboratory Section: New Results.
- [6] Ian C. Bruce, Yousof Erfani, and Muhammad S. A. Zilany. A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites. *Hearing Research*, 360:40–54, March 2018.
- [7] Francis A. M. Manno, Raul R. Cruces, Condon Lau, and Fernando A. Barrios. Uncertain Emotion Discrimination Differences Between Musicians and Non-musicians Is Determined by Fine Structure Association: Hilbert Transform Psychophysics. *Front Neurosci*, 13:902, 2019.
- [8] Christoph Arnoldner, Dominik Riss, Markus Brunner, Martin Durisin, Wolf-Dieter Baumgartner, and Jafar-Sasan Hamzavi. Speech and music perception with the new fine structure speech coding strategy: preliminary results. *Acta Otolaryngol*, 127(12):1298–1303, December 2007.
- [9] Li Xu and Bryan E. Pfingst. Relative importance of temporal envelope and fine structure in lexical-tone perception (L). *J Acoust Soc Am*, 114(6 Pt 1):3024–3027, December 2003.
- [10] Hiroshi Riquimaroux. Perception of noise-vocoded speech sounds: Sentences, words, accents and melodies. *Acoustical Science and Technology*, 27(6):325–331, 2006.
- [11] Xing Li, Kaibao Nie, Nikita S. Imennov, Jong Ho Won, Ward R. Drennan, Jay T. Rubinstein, and Les E. Atlas. Improved perception of speech in noise and Mandarin tones with acoustic simulations of harmonic coding for cochlear implants. *J Acoust Soc Am*, 132(5):3387–3398, November 2012.
- [12] Gavin M. Bidelman and Michael G. Heinz. Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing. *J. Acoust. Soc. Am.*, 130(3):1488–1502, September 2011.