Exploring the Role of Temporal Fine Structure and Envelope in Timbral Coding

Presented By Andrew Sivaprakasam— BME 511 Fall 2020

12/07/2020

Music Coding by the Auditory Periphery

To appreciate an art, we often must make use of one of our senses. But what happens when our senses are impaired? What kind of signals do we receive from our sensory inputs, and in what fidelity do we receive them? What musical features are most important for us to "lock on" to, and what happens when we can't lock on to them?

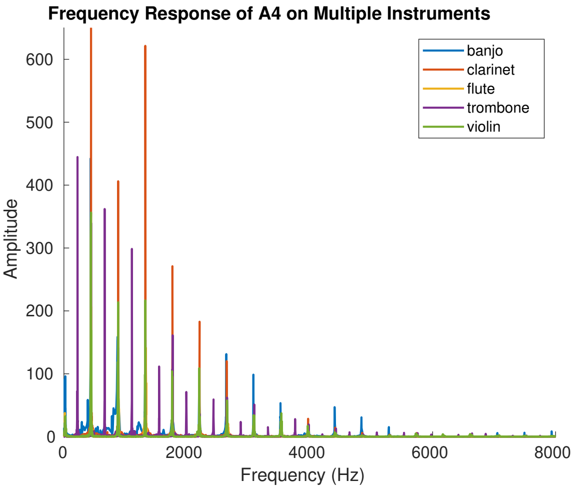
- Musical stimuli are not simple, they are **often non-periodic and spectrally complicated**
- If we look *purely* at how we encode components of music, and how important the fidelity we encode those components is, **we can drive innovative processing algorithms used in hearing aids and cochlear implants to better represent musical sounds**

A start would be looking at how **timbre** is encoded by the auditory system, and what the signals "look like" before being sent to the brain

What is *timbre*?

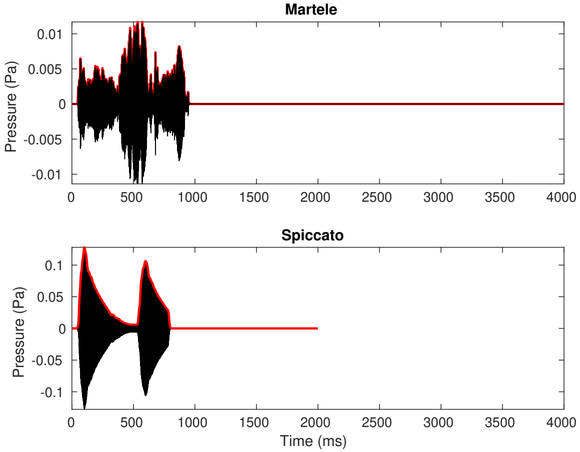
Timbre is the *color* or *quality* of a tone perceived by a listener

- It is a complex psychoacoustic phenomenon that still is not completely understood
- Timbre is what helps listeners differentiate when the same tone is played by a different instrument
- Most consider instrumental timbre to be due to the difference in the magnitude of the harmonics of a given tone



What is *timbre*?

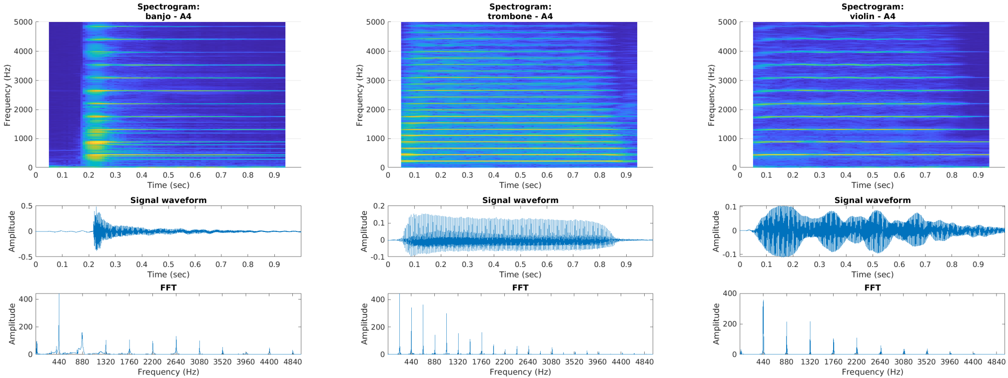
- However, envelope is also important
- In fact, a 2011 study by Heng *et al.* demonstrated that cochlear implant users can differentiate between instruments when 0% TFS is used in a music-noise chimaera
- Normal hearing users, however, rely primarily on TFS cues to differentiate instruments



Envelope cues could help us differentiate between different articulations in music, such as *martele* and *spiccato* played by the violin

Spectrotemporal Analysis

Spectrograms are an useful tool to visualize differences in instruments. Here are a few of the stimuli I used. The sound samples were collected by the Philharmonia Orchestra in London.



The Idea

Goal: Use a spectrally-specific framework designed to inspect the relevance of TFS and ENV in auditory nerve responses to investigate timbral coding in normal hearing and hearing impaired conditions.

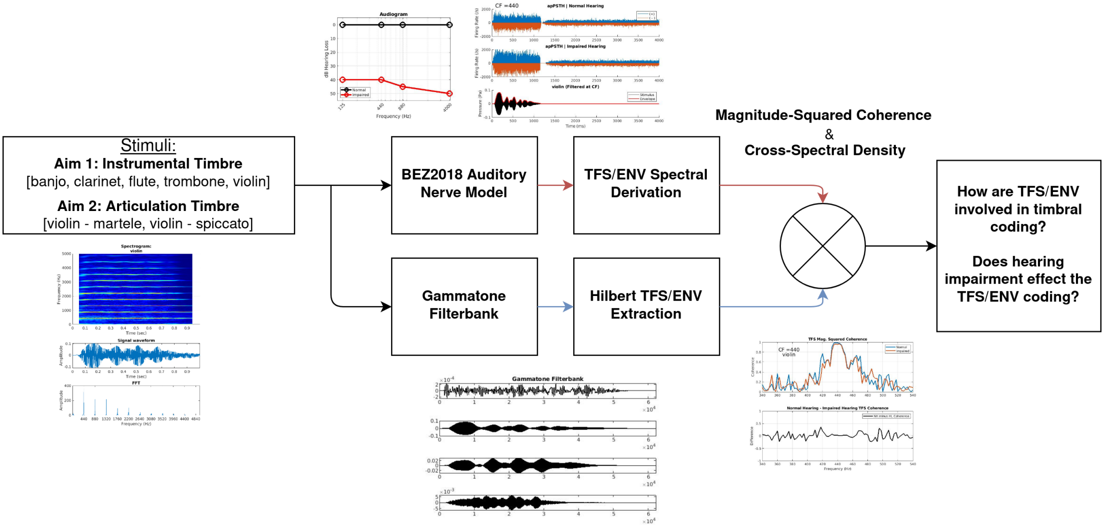
Aim 1: **Instrumental** Timbral Coding:

- Compare timbral coding differences between **banjo, clarinet, flute, trombone, violin** for a common pitch, **A4** (440 Hz)

Aim 2: **Articulation** Timbral Coding:

- Compare timbral coding differences between two articulations on the violin, **martelé** and **spiccato**

General Methods Overview



Modeling Approach

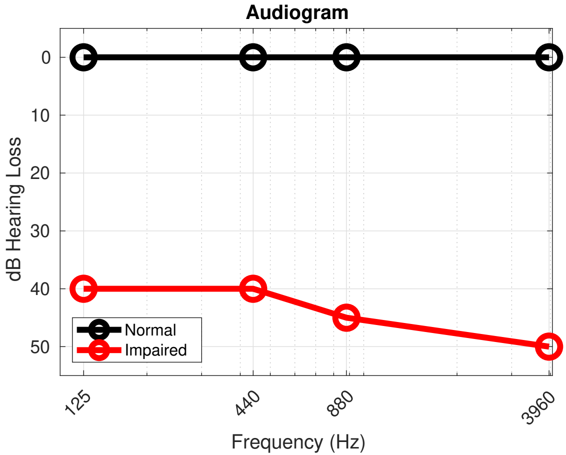
- Using a model of the auditory nerve developed by Zilany et al. and updated by Bruce et al. in 2018, I was able to quickly collect data.
- This model can be run in Matlab and also has a variety of customizable parameters.
- I simulated auditory nerve responses at 4 different Characteristic Frequencies (CF):

$$125 \text{ Hz}, F_0 \text{ (440 Hz)}, F_1 \text{ (880 Hz), and } F_8 \text{ (3960 Hz)}$$

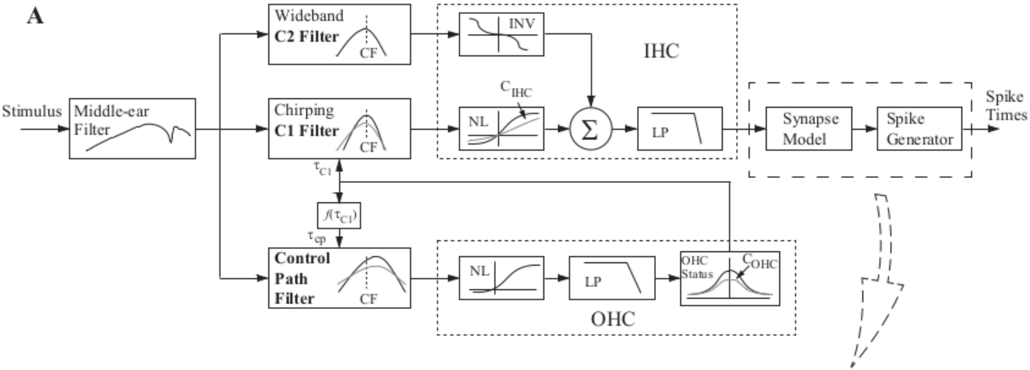
- The Organ of Corti in the cochlea can be thought of as a filterbank with filters centered around a given CF (the further down the cochlea you go, the lower the CF)

Modeling Approach

This model also allowed adjustment of Outer Hair Cell and Inner Hair Cell parameters based on an audiogram. I chose hearing loss threshold shifts roughly based on data reported by Tufts et al. in a 2005 study of musical interval consonance:



Modeling Approach



A schematic of the original Zilany Model. Sound stimulus goes in, spike train comes out.

Model Parameters and Implementation

Using BEZ2018 Model:

Parameters:

- CF (Hz) = 125, 440, 880, 3960
- Repetitions/CF = 75
- Spont. Rate = 70
- Abs and Baseline mean relative refrac. period = 0.6 ms
- IHC/OHC health = set by `fitaudiogram2`
- Species = Human - Shera tuning
- Fractional Gauss. noise type = 0 (fixed)
- Power-law functions = 0 (approximate)
- **Spike time resolution = 10** μs

Other Details:

- Compensated for threshold shift in Hearing Loss condition by increasing stimulus gain by mean threshold shift (43.75 dB)

Model Output