A Repertoire-Agnostic Model of Spectral Fission

Jordan Lenchitz[1]

[1]*College of Music, Florida State University, United States of America*
[1]jlenchitz@fsu.edu

Background

Upper partials of complex tones occupy a peculiar place in many theories of pitch perception. In the case of individual harmonic complexes, the hearing out of partials has long been explored by the likes of Moore, Glasberg, & Peters (1986) through classical psychophysical paradigms that greatly limit the transferability of experimental results to real-life listening situations. As for the perception of simultaneous complex tones (i.e. non-sine wave chords), models of such as that of Terhardt, Stoll, & Seewann (1982) are routinely so concerned with the contribution of spectral-pitch patterns to virtual-pitch percepts that the possibility of conscious perception of spectral components themselves as pitches is often minimized or neglected. A recent theoretical contribution by Lenchitz (2020) provides a framework for understanding prominent spectral components as existing in an auditory attention-modulated border region between timbre and pitch in which certain components may be perceived as pitches of spectral fission, but its computational model relies on assumptions about formant tuning in barbershop performance practice that greatly limit its generalizability. And though existing research by Ladd et al. (2013) has explored individual differences in the context of "spectral listening" and "fundamental listening" for missing fundamental stimuli (i.e. where additively synthesized spectral components are designed to maximize the likelihood of eliciting virtual-pitch percepts), such patterns of individual difference have seldom been experimentally investigated in the cases of simultaneous complex tones that contain energy at fundamentals in addition to featuring prominent upper spectral components.

Aims

To provide a model for analyzing diverse and ecologically valid experimental stimuli with respect to the perception of prominent spectral components as pitches while minimizing *a priori* assumptions about both the audio data and the listener.

Main Contribution

The implementation of a novel spectral prominence detection MATLAB script that takes as input the sound file of any sustained sonority (polyphonic or monophonic) and outputs a two-tiered list of candidate frequencies for spectral fission. After beginning with an overlap-add time stretch with phase randomization to optimize the frequency resolution of the FFT, a noise-tolerant fast peak finding algorithm is applied to the power spectrum, whose results are then narrowed to reflect any dominance of 6 dB or greater by another peak within a 2 ERB radius. These candidates are finally sorted into the two tiers based on whether or not they fall within 2 ERB of any others to minimize the potential role of masking, with those sufficiently separated in the frequency domain from all other candidates placed into the higher tier.

Conclusion and Implications

In light of the results of Alho et al. (2015) showing bottom-up triggered and top-down controlled orienting of auditory attention to pitch activating overlapping brain networks, questions of individual difference with respect to the conscious perception of spectral components as pitches merit investigation from both bottom-up and top-down perspectives. Spectral prominence analysis provides a reference point for both approaches of inquiry into listening behaviors. Preliminary results from a study involving recordings of a cappella singing suggest that this model's predictions of spectral components most likely perceived as pitches of spectral fission are compatible with the self-reports of listeners instructed to listen for high-pitched sounds not being sung by the singers up to difficulties of octave identification in high (1-4 kHz) registers and mediated by individual differences between listeners. The model's two-tiered results can also be used to explore the degree to which spectral fission is robust to stimulus manipulation as well as the influence of learning on the phenomenon in future behavioral experiments. The application of this model to real-world musical examples and additively synthesized stimuli alike provides a quantitative baseline for facilitating the integration of classical psychophysical paradigms with qualitative approaches in the experimental study of spectral fission.

References

Alho, K., et al. (2015). Top-down controlled and bottom-up triggered orienting of auditory attention to pitch activate overlapping brain networks. *Brain Research*, *1626*, 136-145.

Ladd, D. R., et al. (2013). Patterns of Individual Differences in the Perception of Missing-Fundamental Tones. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1386–1397.

Lenchitz, J. (2020, November). Spectral fission in barbershop harmony. Poster presented at the annual meeting of the Society for Music Theory, Minneapolis, MN [virtual].

Moore, B., Glasberg, B., & Peters, R. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *The Journal of the Acoustical Society of America*, *80*, 479-483.

Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *The Journal of the Acoustical Society of America*, *71*, 679-688.

Keywords: spectral fission, computational modeling, individual differences, pitch, timbre.