Farbood, M. M., & Price, K. (2014). Timbral features contributing to perceived auditory and musical tension. In M. K. Song (Ed.). *Proceedings of the 13th International Conference for Music Perception and Cognition*, (pp. 21–26). Seoul, Korea.

TIMBRAL FEATURES CONTRIBUTING TO PERCEIVED AUDITORY AND MUSICAL TENSION

Morwaread M. Farbood

Khen Price

New York University

ABSTRACT

This paper presents two experiments exploring the contribution of timbral features to tension perception. The features examined were inharmonicity, roughness, spectral centroid, spectral deviation, and spectral flatness. The goal of the experiments was to systematically examine how changes in these features contribute directly to changes in perceived tension. In Experiment 1, the stimuli gradually morphed from two extreme states, e.g., "low" inharmonity to "high" inharmonicity. Participants were asked to judge how they felt tension was changing over the course of the sounds. In Experiment 2, the task was to compare static renderings of the low and high states of each feature separated by a silence interval. Listeners were asked to select which of the two sounds was perceived to be more tense. The results of Experiment 1 were complex and inconsistent; the process of morphing introduced certain artifacts that appeared to have influenced listener perception. Furthermore, the process of morphing made it impossible to control for all other features while examining one particular feature. The results of Experiment 2, on the other hand, were very clear—responses clearly indicated that perceived increases in tension correlated to increases in all five features.

1. INTRODUCTION

Timbre is an auditory feature that has received relatively little attention in empirical work examining musical tension. This is in part due to the general difficultly in defining the perceptual dimensions of timbre. While contributions of features such as harmony, loudness, melodic expectation, pitch height, and tempo have been examined thoroughly in past studies, it is still unclear how (and which) timbral features contribute most directly to musical tension. The timbral features that have been examined in past work (from widely varying methodological approaches) include roughness, brightness, spectral flatness, and density (Dean & Bailes, 2010; Helmholtz, 1877; Hutchinson & Knopoff, 1978; Krumhansl, 1996; Nielsen 1987; Plomp & Levelt, 1965; Pressnitzer, McAdams, Winsberg, & Fineberg, 2000).

This study presents two experiments that examine the contribution of five specific timbral features on auditory and musical tension perception: inharmonicity, roughness, spectral centroid, spectral deviation, and spectral flatness. Inharmonicity is a feature that is based on how partials are offset from integer multiples of the fundamental frequency of a pitch. Roughness, described by Plomp and Levelt (1965) as sensory dissonance roughness, is present when pairs of sinusoids are close enough in frequency such that the listener experiences a beating sensation. Spectral centroid is often associated with the perceived brightness of a sound (although the two are not synonymous), and is defined as the geometrical mean of the energy found in the different frequency bins that are produced by a fast Fourier transform (FFT) or any other applicable transformation between the time and frequency domains. Spectral deviation (also termed spectral spread) is attained by calculating the spread of the energy distribution across the spectrum. For example, a pure tone will have no spectral deviation. The spectral flatness (noisiness) of a signal corresponds to how similar its spectrum is to white noise. Low spectral flatness corresponds to more pitched sounds.

The goal of this study was to systematically examine how changes in these features contribute directly to changes in perceived tension. The "directionality" of features was also examined; that is, whether increases in the features contribute to increases in tension, or if this relationship is not clear.

2. GENERAL METHOD

For both experiments, two states were generated for each feature: one with a minimum of that particular character (state A) one with a maximum of that particular characteristic (state B). Since loudness is known to be a dominant factor in auditory perception experiments, loudness was equalized for all stimuli. When possible, an effort was made to avoid covariance between features when synthesizing the stimuli, with special attention given to keeping spectral centroid constant (in the cases where it was not the targeted feature). This is due to the established importance of this feature in timbre discrimination experiments.

Overall intensity level equalization was automatically done using the Echonest API (Jehan, 2010) so that loudness differences across stimuli were less than 1 dB. The Genesis Loudness Toolbox (Genesis, 2009) for Matlab was used to obtain time-dependent loudness measurements for each synthesized timbre for verification purposes. Amplitude envelopes and changes in intensity were applied to compensate for temporal changes in loudness, as well as differences in loudness between stimuli. The timbral features were analyzed using the MIRtoolbox (Lartillot, Toiviainen, & Eerola, 2008). Filters and noise used in the generation of spectral flatness stimuli were created using the Matlab DSP toolbox.

3. EXPERIMENT 1

3.1 Method

For each feature, a stimulus was created that morphed from state A to B, B to A, or did not change over time. In addition there were four different durations for each stimulus: 1, 5, 10, and 15 secs, resulting in a total of 60 stimuli (5 features x 3 change types x 4 durations). Fifty subjects, mostly musicians, took part in the study. The stimuli were presented to listeners over headphones in random order. Participants were asked to judge how they felt tension changed over the course of each sound by selecting one of the following multiple choice responses: increasing, decreasing, no change, or none of the above.

3.2 Results

The results of Experiment 1 were inconsistent and contradictory. There was few clear trends or "directionality" indicated in the responses. It appeared that the process of morphing introduced certain artifacts such as unintended pitch glides that appeared to influenced listener perception; this was particularly a problem for inharmonicity. Increase in inharmonicity was most strongly correlated with increase in tension when the stimuli duration was longer, most likely because the pitch shifts became slower and less noticeable.

Both pitch and loudness appeared to be an issue for spectral flatness due to the perception of crescendos and decrescendos resulting from the source separation of noise and pitch (perceived fading or increase in one or the other). Depending on which aspect (noise vs. tone) the listener's attention was drawn to, either an increase or decrease in loudness could have been salient. This would explain the conflicting results.

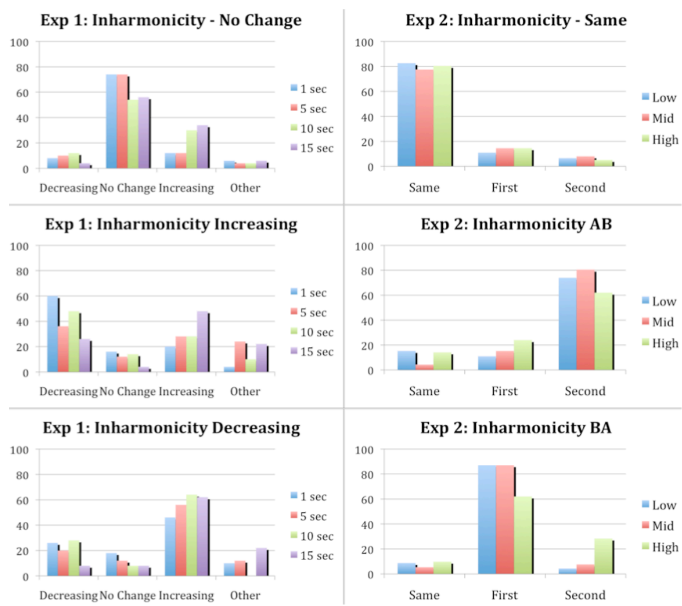

Figure 1: Response profiles for stimuli featuring changing inharmonicity; graphs in the left column show results from Experiment 1, graphs in the right column from Experiment 2.

A perceived change in loudness and/or pitch might have also affected the results for spectral centroid, although from a qualitative perspective, pitch was not distinctly tonal in the stimuli. Spectral centroid had the most scattered and difficult to interpret results.

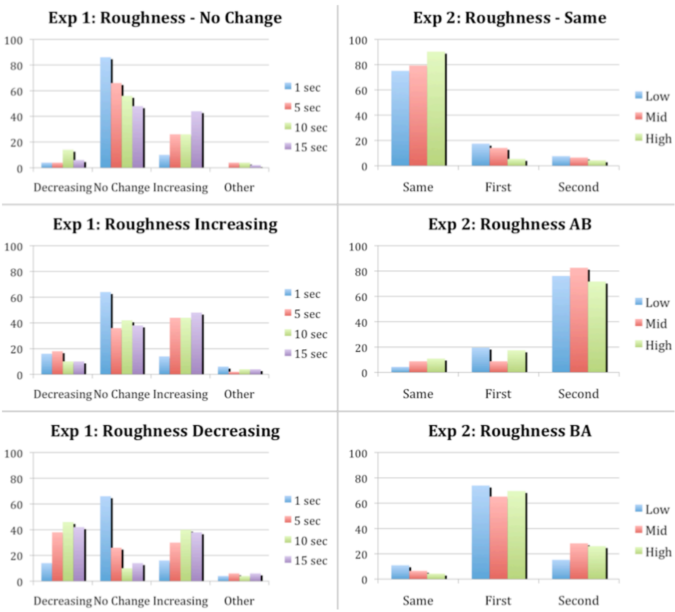

Figure 2: Response profiles for stimuli featuring changing roughness; graphs in the left column show results from Experiment 1, graphs in the right column from Experiment 2.

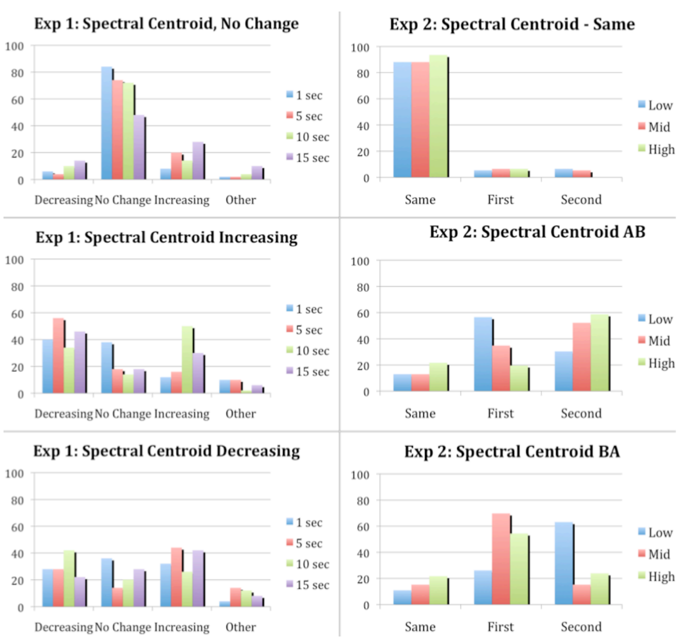

Figure 3: Response profiles for stimuli featuring changing spectral centroid; graphs in the left column show results from Experiment 1, graphs in the right column from Experiment 2.

In the case of stimuli of longer durations, the unchanging versions were not necessarily static in loudness from a psychoacoustic perspective—having a constant sound does not necessarily translate to static perceived loudness, especially if the sound is unpleasant and tense to begin with. This would partially explain the results for spectral deviation, where tension increases were frequently indicated by listeners despite the fact that the sound remained unchanged.

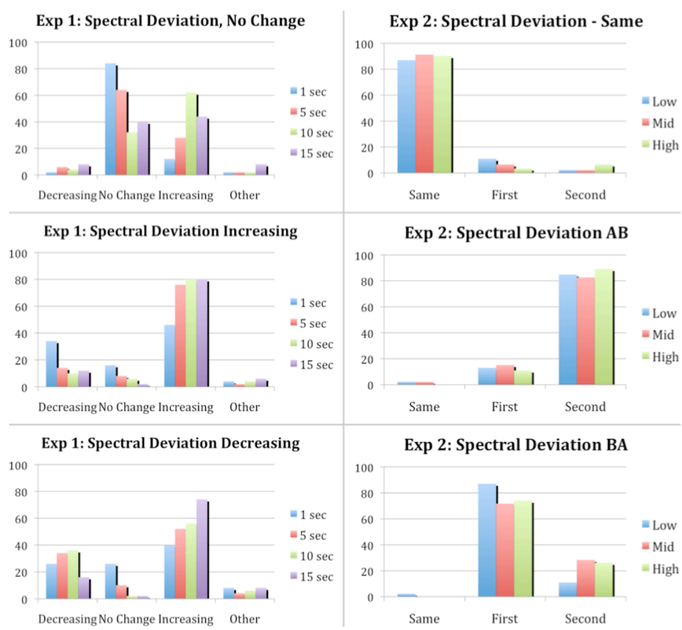

Figure 4: Response profiles for stimuli featuring changing spectral deviation; graphs in the left column show results from Experiment 1, graphs in the right column from Experiment 2.

Although spectral centroid was held relatively constant for stimuli focusing on the other four features, those other features were not completely controlled for. This might have further complicated the results. See Figures 1-5 (left column) for graphs showing the response profiles. Given these mixed results, a second experiment was designed to try to avoid the pitfalls encountered in morphing sound from one state to another.

4. EXPERIMENT 2

4.1 Method

Experiment 2 featured stimuli designed to compare static versions of states A and B separated by 1.2 sec of silence. In addition, there were three different pitch registers utilized for each condition (low, mid, high). For all A/B state pairs there were four possible orderings: AB, BA, AA, and BB.

Forty-six subjects, mostly musicians, took part in the study. The stimuli were presented to listeners in random order over headphones. Participants were asked to judge which of the two sounds (first or second) was more tense or if they sounded the same.

4.2 Results

Unlike the case for Experiment 1, the results of Experiment 2 were very consistent. Increases in all the features corresponded strongly with changes in tension (Figures 1-5, right column). The only interaction effect with pitch register occurred in the case of spectral centroid. Those results indicated that the higher the pitch register, the stronger the correlation between increases in spectral centroid and tension.

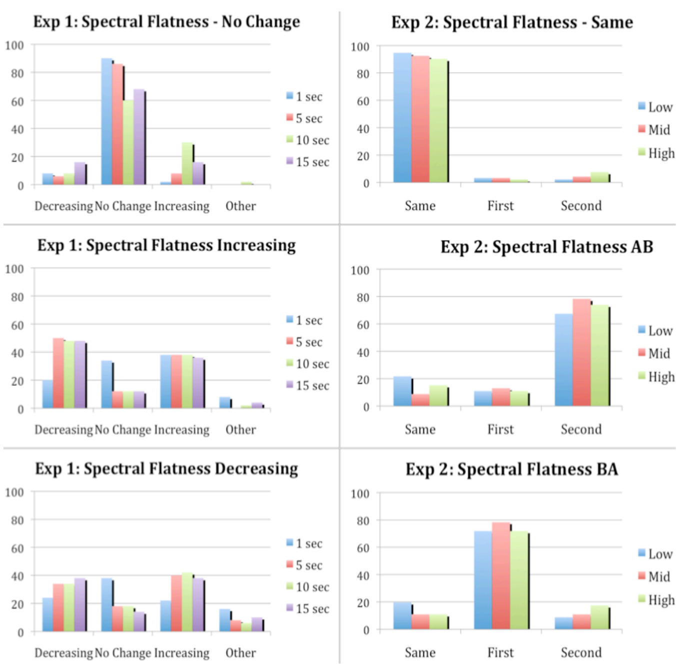

Figure 5: Response profiles for stimuli featuring changing spectral flatness; graphs in the left column show results from Experiment 1, graphs in the right column from Experiment 2.

5. CONCLUSIONS

The five timbral features examined in this study––inharmonicity, roughness, spectral centroid, spectral deviation, and spectral flatness—were shown to contribute to perceived tension. However, changes in some of these parameters were difficult to isolate because of feature covariance. As such, these conclusions were only apparent in the results of Experiment 2. The results of Experiment 1 were complex and inconsistent due to the inability to control for all features as well as pitch and loudness. Experiment 2 used static versions of the two extreme states for each feature, resulting in much clearer response profiles. In all cases, an increase in each feature corresponded to an increase in perceived tension.

The next step is to explore precisely how these features covary in order to model how dynamic timbral changes influence tension perception. Additional experiments using more complex stimuli—particularly musical stimuli where other musical features influencing tension such as harmony and melodic contour are involved—are the next directions to explore in future work.

6. REFERENCES

Dean, R. T., & Bailes, F. (2010). Time series analysis as a method to examine acoustical influences on real-time perception of music. *Empirical Musicology Review*, *5*(4), 152–175.

Gensis (2009). Loudness Toolbox. Retrieved from http://www.genesis-acoustics.com /en/index.php?page=32

Helmholtz, von, H. L. F. (1877). Die Lehre von der Tonempfindungen als physiologische Grundlage für die Theorie der Musik. New York: Dover.

Jehan, T. (2010). Echo Nest Remix API. Retrieved from http://developer.echonest. com/client_libraries.html

Krumhansl, C. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception, 13*(3), 401–432.

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab toolbox for music information retrieval. In *Studies in Classification, Data Analysis, and Knowledge Organization: Data Analysis, Machine Learning and Applications* (pp. 261–268). Springer.

Nielsen, F. (1987). Musical "tension" and related concepts. *The Semiotic Web '86: an International Yearbook*, 491–514.

Plomp, R., & Levelt, W. J. M. (1965). Tonal Consonance and Critical Bandwidth. *Journal of the Acoustical Society of America*, *38*, 548–560.

Pressnitzer, D., McAdams, S., Winsberg, S., & Fineberg, J. (2000). Perception of music tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Perception & Psychophysics*, *62*(1), 66–80.