Confounding Pitch Height in Studies of Tonal Hierarchy

V. J. Manzo
Boyer College of Music
Temple University
Philadelphia, PA 19122
vjm@temple.edu

ABSTRACT

In studies of perception of harmonic direction, pitch height is among the factors that must be confounded by researchers. Since the 1980's, several differing techniques have been used. This paper explains the techniques used, exposes several shortcomings, and provides improvements to these models developed by the author. These models are available for download as open-source and standalone cross-platform applications from the author's website at http://www.vjmanzo.org.

Categories and Subject Descriptors
General Terms: Perception of Harmony

Keywords: Models of Harmonic Perception, Music Theory, Harmonic Direction, Pitch height, Shepard Tones, Tonal hierarchy

1. INTRODUCTION

"Arguably, listeners must establish the key of a piece within the first few bars (if not notes) if they are to code and store musical material appropriately "(Sloboda, 2005, p. 129). In many seminal studies of the perceptual organization of harmony, researchers have used similar models that avoid complications associated with pitch height while conducting experiments. These models are based on perception and are designed to confound attributes of pitch height in harmony such as voice leading and inversion leaving only pure harmony for the participant to hear and thereby make judgments (Deutsch & Boulanger, 1984).

Early studies in perceptual organization of harmony implemented models to confound pitch height based on the so-called *Shepard Tones* named after Roger Shepard. Shepard is credited with creating a model similar to Escher's ever-ascending staircase in which the listener perceives a pitch to be continually rising in an infinite glissando (Shepard, 1964). In *Shepard Tones,* the only harmonics are multiples of the fundamental frequency that are powers of 2. That is, f_0, $2f_0$, $4f_0$, and $8f_0$ may be present, but $3f_0$ and $5f_0$ cannot. This is expressed as:

$$y(t) = \sum_i \cos\left(2\pi\left(f_0 \cdot 2^i\right)t\right)$$

A Shepard scale comprises every multiple of a single frequency at each audible octave spanning the full range of hearing (20Hz – 20 kHz). This produces an overtone series unlike any orchestral instrument where the overtone series would normally comprise tones built from multiplying consecutive integers by the fundamental frequency. That is, f_0, $2f_0$, $3f_0$, $4f_0$, etc.

Playing each sine wave in the *Shepard Tones* simultaneously at equal amplitude levels sounds similar to the timbre of an organ, though it does not have the same overtone series. A fixed Gaussian filter enhances the illusion that a pitch is continually rising. Higher frequencies appear to taper off to the right side of the filter as lower frequencies appear to grow thus giving the illusion of infinite glissando.

2. MODELS IN HARMONIC PERCEPTION

Shepard's research has had considerable influence on the models used in experiments of harmonic perception. Richard Moore's implementation of this phenomenon served as the basis of harmony used in the research of Deutsch, Moore, and Dolson (1984) with some variation. Shepard's model created an illusion that unraveled over time. Using the model for experiments in perception should account for all tones across the audible spectrum happening simultaneously instead of linearly as in Shepard's research.

Moore's implementation used a stationary Gaussian filter on the chord stimuli, three tones built for each chord spanning a seven octave bandwidth. The outer two octaves, however, had an amplitude of zero as a result of the Gaussian bell curve. This stationary Gaussian filter, as used in Shepard's model, did not continue filtering the same bands relative to a different pitch class. For example, filtering pitch class C0 would result in a 0 zero amplitude at frequency 0 and frequency 6, whereas filtering pitch class D0 would result in a non-zero amplitude at frequency 0 and a zero amplitude at frequency 5 and frequency 6.

The notion of a stationary Gaussian filter is an artifact of the *Shepard Tones* and could be problematic for confounding pitch height since a stationary filter will bias the output toward a center frequency. A moveable Gaussian filter centered around a specific frequency ($4f_0$ for example) would allow the model to maintain equal filtered relationships while changing pitch. The current model also disregards the concept of equal loudness contours (Fletcher & Munson, 1933). The frequencies that are being tapered off at each end of the spectrum are the ones that humans generally have more difficulty hearing, so it would seem that if hearing all pitches at an equally perceived level will confound pitch height, the curve should more closely resemble a peaknotch filter than a Gaussian filter.

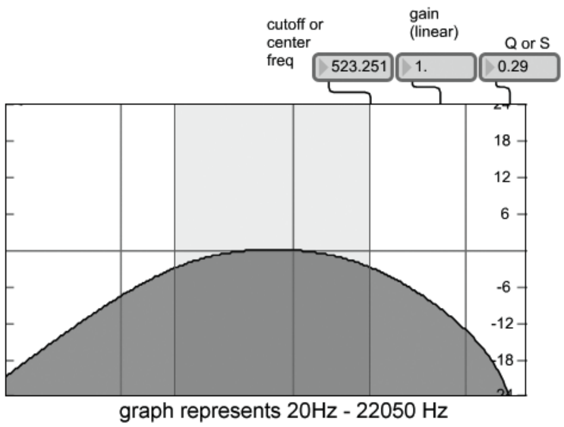

Figure 1. A bandpass filter centered at 64f_0 where f_0 = C0

Krumhansel, Bharucha, and Kessler (1982) had implemented a similar model two years earlier that controlled the amplitude of the chord stimuli, three tones built for each chord spanning a five octave range, with a bell curve where its outer bands tapered off. The tones comprising each chord were filtered with a Fletcher-Munson equal loudness curve. It is unclear if the bell curve was stationary or moveable with respect to the center frequency (a multiple of the fundamental).

The researchers state, “The chords were constructed to obscure the topmost and bottommost tones of each triad.” They claim that this method “…produces chords without a clearly defined highest component, minimizing melodic effects”. However, the center two octaves were at the same level while the other octaves tapered off which, despite the Fletcher-Munson curve, may bias the listener toward a central frequency or a group of central frequencies instead of yielding a sound where all audible frequencies are present at a perceived equal loudness.

They continue, “An alternative way of eliminating pitch height differences in the present study might have been to select the three tones of the triads from different octaves, yielding different "inversions" so that there would be no large overall pitch height differences between chords. This method of controlling for pitch height, however, would be complex and only approximate. In addition, different inversions of the same triad sound slightly different and do not play identical roles in musical composition.”

Other studies related to tonal hierarchies have used similar models with slight variation such as Krumhansl and Kessler (1982). In this study, each chord tone (dominant seventh chords were also used in this experiment in addition to triads) spanned a five octave range, with a bell curve where its outer bands tapered off.

The above three studies all attribute the technique of their model to Roger Shepard though, they all differ from his original model and from each other.

Without the presence of all audible multiples of the octave, these models may not confound the effects of pitch height such as inversion. Perhaps the ideal model in a Shepard style chord generator would be to have all frequencies at a perceived equal loudness using an equal loudness contour.

One problem with quantifying equal loudness contours is that they are based on perception research: there is no absolute equal loudness contour only sets of contours that conform to different age demographics. The human hearing system is more sensitive to some frequencies than others, so some frequencies are perceived as having a different loudness than others even if they have the same amplitude. Sound pressure levels (dB SPL) are measured in phons over the frequency spectrum. As phons increase or decrease, the perceived loudness of frequencies also changes, thus equal loudness contours change in response to the number of phons.

For about 23 years, the Fletcher-Munson equal loudness curve was the standard contour for representing frequencies at a perceived equal loudness. In 1956, the Robinson-Dadson contour (Robinson & Dadson, 1956) became the standard and remained so until 2003. Currently, the International Organization for Standardization created the ISO226:2003 contour (International Organization for Standardization) based on the Robinson-Dadson which is the current recognized standard.

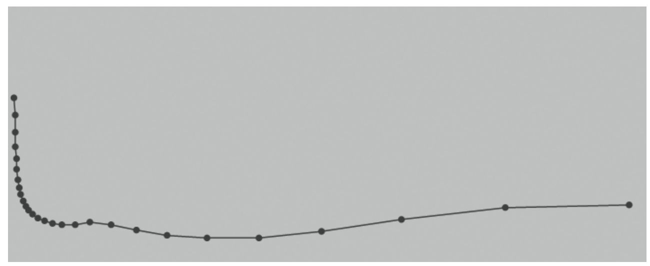

Figure 2. An ISO226:2003 contour at 1 phon

There is an apparent need to replicate the above mentioned studies as well as other extant studies related to tonal hierarchies incorporating the current ISO226:2003 contour where the Fletcher-Munson contour was used in order to strengthen or refute previous finding based on older chord generation models. Studies that did not incorporate an equal loudness contour such as Deutsch, Moore, and Dolson (1984) should also be replicated.

The practical application of each of the research models should also be reconsidered with regard to the overtone series that is being tested. As earlier stated, no orchestral instrument bears the overtone series construction formed by multiplying the lowest audible fundamental frequency by multiples of two. Resultant data may not be transferable in contexts outside of those in which an instrument of this timbre is used.

Perhaps a model that resembles more closely a larger range of extant musical instruments should be developed in place of a model that resembles practically none. Perhaps there is data that exists in the harmonic series of different instruments that serve as indicators for establishing tonal hierarchies. The research of Paraskeva and McAdams (Paraskeva & McAdams, 1997) suggests that some timbres may give the listener different cues for analyzing and establishing tonal hierarchies than others.

3. NEW COMPUTER MODELS

I have created a computer model of the original tone generation models used by Krumhansel, Bharucha, and Kessler (1982) and have given it the capability to allow the filter to span any number of octaves as in the research of Deutsch, Moore, and Dolson (1984). I have also included a switch to allow the center frequency of the filter to shift with each new fundamental frequency causing each overtone series to receive uniform filtering. This model also allows the user to change the Gaussian filter to an allpass filter or a peaknotch filter.

I have also created a computer model that departs from the Shepard-based tone generation models and implements the ISO226:2003 standard. The user will have to ensure that their loudspeakers are measures the same phon level as selected in this model.

4. CONCLUSION

With advances in new technology, it is prudent to revisit seminal studies in perception of tonal hierarchies and apply new models to support prior research. The models discussed herein are available for download from the author's website http://www.vjmanzo.org or http://www.vjmanzo.com/cv.

5. REFERENCES

[1] Deutsch, D., & Boulanger, R. C. (1984). Octave Equivalence and the immediate recall of pitch sequences. *Music Perception , 2*, 40-51.

[2] Deutsch, D., Moore, F. R., & Dolson, M. (1984). Pitch classes differ with respect to height. *Music Perception , 2*, 265-271.

[3] Fletcher, H., & Munson, W. A. (1933). Loudness, Its Definition, Measurement and Calculation. *Journal of the Acoustical Society of America , 5*, 82-108.

[4] International Organization for Standardization. (n.d.). *ISO 226:2003*. Retrieved March 10, 2009, from International Organization for Standardization: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34222

[5] Krumhansel, C. L., Bharucha, J. J., & Kessler, E. J. (1982). Perceived harmonic structure of chords in three closely related musical keys. *Journal of Experimental Psychology: Human Perception and Performance , 8*, 24-36.

[6] Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the Dynamic Changes in Perceived Tonal Organization. *Psychological Review , 89*, 334-368.

[7] Paraskeva, S., & McAdams, S. (1997). Influence of Timbre, Presence/absence of Tonal Hierarchy and Musical Training on the Perception of Musical Tension and Relaxation Schemas. *CMC: International Computer Music Conference* (pp. p.438-441). Thessaloniki: Ircam - Centre Georges-Pompidou.

[8] Robinson, D. W., & Dadson, R. S. (1956). A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics* , 166-181.

[9] Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America , 36*, 2346-2353.

[10] Sloboda, J. (2005). *Exploring the musical mind.* New York: Oxford University Press Inc.