

Commentary on Keith Mashinter’s “Calculating Sensory Dissonance: Some Discrepancies Arising from the Models of Kameoka & Kuriyagawa, and Hutchinson & Knopff”

RICHARD PARNCUTT
University of Graz

ABSTRACT: Mashinter’s (2006) mathematical model of sensory dissonance neglects the dependence of roughness on waveform, the role of masking, the distribution of roughness across critical bands, the possible positive contribution of fusion or toneness to euphony, and the familiarity and music-theoretical functions of a sonority. Of course not all these aspects can reasonably be included in a model, but they can affect the data with which its predictions are compared.

Submitted 2006 August 12; accepted 2006 August 15.

KEYWORDS: *consonance, dissonance, perception, sensory, enculturation*

A fundamental problem in models of the perception of musical structure is the old question of how to deal with the interaction between sensory (bottom-up, primitive, innate) and cultural (top-down, schema-based, learned) aspects (e.g., Bregman, 1990; Parncutt, 1989; Parncutt & Bregman, 2000). The boundary between sensory and cultural is not clearly defined and depends on its (implied or explicit) operational definition. One possible interpretation is that the sensory or bottom-up aspect corresponds to the origin of the phenomenon, while the cultural or top-down aspect to its empirically observable musical manifestation.

The relationship between sensory and cultural is of course central for any study of musical dissonance. Previous studies on musical dissonance have tended to focus on one or the other of these two aspects. Mashinter’s study (2006) continues that tradition by focusing on the sensory aspect. I argue in the present commentary that these two aspects of dissonance cannot be clearly distinguished.

A sonority may be judged *dissonant* for a number of reasons. In western music, the most important of these are presumably how *rough* it sounds (i.e. whether the sound seems to have a rough surface – the sensory, bottom-up, primitive, innate aspect) and how *familiar* it is, or how often the listener has been exposed to it in music (the cultural, top-down, schema-based, learned aspect). The latter in turn depends on such scientifically slippery factors as the kind and clarity of structural (harmonic, melodic) functions that the sound may fulfill in music. But there may also be a positive aspect – an independent component of *consonance* - that scientists have so far failed to nail down. The aesthetic question of why consonant polyphony is beautiful has never adequately been answered: I am not aware of any clear psychological reason why fusion in the sense of Stumpf, in which the perceived number of tones is less than the actual number, should make a distinct positive contribution to overall consonance.

To illustrate the complex interaction between the various factors that may influence consonance and dissonance, consider the example of a major-seventh chord (e.g. from C3-G3-B3-E4). Why does this chord sound more consonant than an isolated major-seventh interval (C3-B3)?

- Is it because the beating harmonics within the dyad C3-B3 are partially masked when the tones G3 and E4 are added? If so, which model of roughness accounts for this?
- Is it because the roughness in the chord is more evenly distributed across different critical bands than in the dyad – which, according to Plomp and Levelt (1965), also explains why the interval between tenor and bass voices in typical musical chords tends to be bigger than the interval between soprano and alto?
- Is it because the chord has a clearer function in music, and is therefore more familiar, than the dyad?

- Is it because the chord's fusion is associated with an additional, independent aspect of "euphony"?

All psychoacoustical data on the dissonance (or roughness) of musical sounds are potentially "contaminated" with the musical experience and preferences of listeners. This problem can ultimately never be solved, but we could at least get more insight into it by creating a database of dissonance judgments of a large set of musical and non-musical sounds by a wide range of western and non-western musicians and non-musicians, including a considerable number of people who have very little experience of western music. This project is of course methodologically and ethically problematic, because different musical cultures have different concepts of dissonance and it is difficult to express the western concept of dissonance in many non-western languages. One might at least ask listeners how much they like or dislike a given sound and hope that, in a large enough database, differences in cultural associations, optimal complexity and so on will somehow cancel each other out. In any case, this project remains to be carried out (Are there any travel-hungry students out there looking for a Ph.D. topic in the area of ethno-psycho-musicology?) and is at least part of the reason why no-one has yet been able to create a reliable and widely accepted model of the roughness of musical sounds. One would first need to create such a data set and then to compare different models, and different variants of each model, with the data.

An easier problem to solve surrounds the definition and use of the terms sensory dissonance and roughness. The main authors cited by Mashinter – Kameoka and Kuriyagawa (1969), Plomp and Levelt (1965), and Hutchinson and Knopoff (1978) – confuse these two terms when they speak about "sensory dissonance" but conceive or model it in terms only of roughness based on rapid beating. If sensory dissonance is the sensory component of musical dissonance, it depends on several factors. According to Terhardt (see summary at <http://www.mmk.ei.tum.de/persons/ter/top/senscons.html>), these include not only roughness but also sharpness, loudness, and toneness (*Klanghaftigkeit*, or the extent to which a sound evokes sensations of tone rather than noise). Of these, toneness is the only positive factor, and may be related to the idea of euphony referred to above – the beauty of consonant polyphony – via harmonic or virtual pitch salience. Mashinter's discussion could be clarified by adopting Terhardt's terminology and talking about roughness rather than sensory dissonance.

That existing models of roughness do not work very well became clear to me as I attempted to introduce roughness into a quantitative explanation of the tension judgments of regular musical chords (sonorities of harmonic complex tones) reported in Bigand, Parncutt, and Lerdahl (1996). Calculations according to my implementation of Hutchinson and Knopoff (1978) (<http://www-gewi.uni-graz.at/staff/parncutt/computerprograms.html>), which includes reasonable assumptions about the amplitude envelope of each tone, didn't correlate very well with the tension data. The correlation improved when we simply counted the number of semitones, tritones, major sevenths and their compounds in each sonority. This experience was consistent with the idea that the top-down or cultural aspect of the dissonance of typical sonorities may be as strong as, if not stronger than, the bottom-up or sensory aspect.

Further possible reasons why the cited models fail are as follows.

- They are confined to the frequency domain. If roughness really does correspond to the bumpiness of the acoustic surface of a sound, it should be modelled in the time domain (Aures, 1985). That phase effects play an important role was demonstrated by Pressnitzer and McAdams (1999).
- The result depends on the formula for critical bandwidth, or the dependency of the roughness of two pure tones on their frequency interval. Modern modelers can choose among various possible formulae (Greenwood, 1961; Moore & Glasberg, 1983; Plomp & Levelt, 1965; Rakowski, 1982; Zwicker, 1982). When optimizing model performance, it should not be taken for granted that roughness as a function of frequency interval is directly related to critical bandwidth, however operationally defined. Rather, one should search for the mathematical formulation that yields predictions that best match a wide range of empirical data.
- The models do not account for the possible role of critical bands in the summation of different contributions to overall loudness (Terhardt, 1974; Aures, 1985). Roughness perception is presumably similar to loudness perception in that contributions from different critical bands add more strongly than contributions from within a single critical band.

- If one is working in the frequency domain, it is not necessarily true that roughness contributions add linearly. Kameoka and Kuriyagawa (1969) assumed that contributions should be raised to the power 4, added, and the sum raised to the power 0.25. But these numbers, or this approach, can presumably be improved.

Given these numerous problems, which in spite of a century of research on sensory dissonance have still not been satisfactorily solved, I consider it premature to comment on Mashinter's mathematical formulations. The proof of any new model of roughness will be in the pudding: the model must first successfully account for a wide variety of empirical data, and the appropriateness of its mathematical formulations should primarily be evaluated on that basis.

REFERENCES

- Aures, W. (1985). Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale. *Acustica*, Vol. 59, pp. 130-141.
- Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: Influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, Vol. 58, pp. 125-141.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: MIT Press.
- Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America*, Vol. 33, pp.1344-1356.
- Hutchinson, W. & Knopoff, L. (1978). The acoustical component of western consonance. *Interface*, Vol. 7, pp. 1-29.
- Kameoka, A. & Kuriyagawa, M. (1969). Consonance theory. *Journal of the Acoustical Society of America*, Vol. 45, pp. 1451-1469.
- Mashinter, K. (2006). Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. *Empirical Musicology Review*, Vol. 1, No. 2, pp. 65-84.
- Moore, B. C. J., & Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America*, Vol. 74, pp. 750-753.
- Parncutt, R., & Bregman, A. S. (2000). Tone profiles following short chord progressions: Top-down or bottom-up? *Music Perception*, Vol. 18, pp. 25-57.
- Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, Vol. 38, pp. 548-560.
- Pressnitzer, D., & McAdams, S. (1999). Two phase effects in roughness perception. *Journal of the Acoustical Society of America*, Vol. 105, pp. 2773-2782.
- Rakowski, A. (1982). Psychoacoustic dissonance in pure-tone intervals: Disparities and common findings. In Dahlhaus, C. & Krause, M. (Eds.), *Tiefenstruktur der Musik* (pp. 51-67). Berlin: Technische Universität Berlin.
- Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). *Acustica*, Vol. 30, pp. 201-213.
- Zwicker, E. (1982). *Psychoakustik*. Berlin: Springer.