

A Response to Cross & Rohrmeier's 'Comments on Facilitation and Coherence Between the Dynamic and Retrospective Perception of Segmentation in Computer-Generated Music'

FREYA BAILES

MARCS Auditory Laboratories, University of Western Sydney

ROGER T. DEAN

MARCS Auditory Laboratories, University of Western Sydney

ABSTRACT: The commentary by Cross and Rohrmeier (2007) attempts to locate our paper (Bailes and Dean, 2007a) as a study of timbre, and points out the ongoing development of research in this area, including attempts to define psychoacoustic thresholds of perception. However, our work is directed to understanding broader psychological phenomena such as the impact of sound duration on the perception of structure in computer music, and the concordance between real-time and retrospective measures. We discuss further our identification of an asymmetrical detection of sound segmentation, questioning the conceptual distinctions of timbre perception that Cross and Rohrmeier propose.

Submitted 2007 September 28; accepted 2007 September 28.

KEYWORDS: *music perception, segmentation, cognitive discrimination, computer music*

IN commenting on our working paper, 'Facilitation and coherence between the dynamic and retrospective perception of segmentation in computer-generated music', Ian Cross and Martin Rohrmeier (2007) acknowledge that it addresses a neglected area of research, namely the relationship between human perception and contemporary music composition. However, their comments focus almost exclusively on the question of timbre perception, distracting from the other, equally fundamental aspects of the research we report. Timbre perception is only part of what occurs when a listener is asked to discriminate between a segmented or non-segmented sound sequence. In particular, our study highlights broader psychological phenomena for which there is increasing evidence of their operation across different musical dimensions. We find coherence between the responses of non-expert listeners in real-time and in retrospect. We also observe a 'ramp archetype' for computer-generated textures, drawing a parallel with auditory looming (initially understood as a spatial effect). Similarly, we argue for the consideration of sound duration and temporal order to account for context effects, which have seemingly little to do with timbre properties of the stimuli we study. Nevertheless, refining concepts of timbre would certainly be of significant theoretical interest, as researchers finally begin to empirically address listener perceptions of electroacoustic composition.

The definition of timbre is notoriously problematic, even for music performed on traditional acoustic instruments. Cross and Rohrmeier (2007) suggest that contemporary understandings of the nature of timbre perception are divided into three distinct conceptual approaches: 1) timbre experience derives from our innate understanding of the relationship between acoustical attributes of a sound signal and properties of the mechanical system producing the sound; 2) our understanding of the relationship between sound signal and sound production is learned; 3) listener experience of timbre correlates with the "abstract characteristics of the sound structure" (p. 147). We have argued elsewhere (Bailes and Dean, 2007b) for the importance of examining the intersection of listener perceptions of the relationship between sound properties and sound production, alongside more abstract sound characteristics not readily related to a sound-generating event. Evidently, computer-generated composition is peculiar in its ability to defy the

usual rules of mechanical sound production. Van Nort (2007) argues that timbral properties of electroacoustic composition that do not evoke a sound-generating action might be more usefully described by the concept of sonic texture than timbre.

Cross and Rohrmeier (2007) point out that our finding of a ramp archetype in the perception of computer-generated music adds weight to theories of a perception of timbre shaped by our understanding of environmental regularities. The ramp archetype describes the phenomenon in which listeners are better able to perceive an increase in sound intensity than a decrease of the same magnitude in operation over a similar space of time. Such perceptual asymmetries are seemingly adaptive (Neuhoff, 2001), and probably relate more to a bottom-up perception of sound (albeit mediated by attention) than the higher cognitive processing involved in recognizing sounds. The strategy which listeners adopt when listening to such music is far from clear, and surely depends upon the listening context, personal experience, and the sounds themselves. Electroacoustic composers deliberately harness the potential for timbral ambiguity inherent in the medium (Brattico & Sassanelli, 2000). Along these lines, we have recently conducted research (in preparation) exploring the affective correlates of sonic ambiguity, finding a relationship between the affective measure of sounds and the recognition or otherwise of their derivation. It seems that the conceptual categories of timbre perception outlined by Cross and Rohrmeier are neither mutually exclusive nor comprehensive. Explicating conceptual levels is important to theories of timbre more broadly, but not the purpose of the experiment we report.

A clear advantage to the study of computer-generated music is the degree of algorithmic control afforded when composing the stimuli, without sacrificing ecological validity. However, while a control of the parameters of interest is important, so is our research question that seeks to learn about the perception of structure in music. We place our study firmly within the tradition of music research, and describe our stimuli in at least as much detail as a verbal description of a recording. Musicological research without the visual crutch of a score needs to find alternative ways to describe its subject. Here we are helped by the provision of the specific algorithmic composition of our stimuli, and while Cross and Rohrmeier (2007) criticize a lack of specifiable variation between our sound segments, Table 2 (Bailes and Dean, 2007a, p. 76) is one example of significant detail for the varieties of our 'Forbidden Planet' series of sounds. It is challenging to strike a balance between the provision of a comprehensible verbal description and a technical quantitative analysis of musical sounds. When reporting significant effects relating to the sounds themselves, we are detailed in our description of relevant aspects of their composition, such as the spectral power of our stimulus FB09FB10, before and after the change in segment. Ultimately, there is no substitute for the sound itself, and we provide a number of sound examples on-line within the text. All the stimuli are available from the authors, and have been generated by algorithms that can be completely defined.

Finally, we wish to address the methodological complaint of Cross and Rohrmeier (2007) that we do not attempt to explore detection thresholds via the exhaustive variation of sound parameters in our stimuli. Such a method would certainly be appropriate in a psychoacoustic study of 'just noticeable differences' (JNDs), but is less relevant to the concerns of an investigation of music perception (as discussed in Bailes and Dean, 2005, 2007b). The condition of listening to a continuous sound and deciding whether there are multiple segments within it provides completely different demands from those of psychoacoustic experiments comparing two separate sounds. Similarly, the inclusion of real-time continuous response is quite different from the methods used in retrospective JND studies. Rather, our study emphasizes the differing structural inferences that listeners make during three repetitions of each stimulus, a task that is necessarily focused on cognitive and musical as well as perceptual response.

REFERENCES

Bailes, F., & Dean, R. T. (2005). Structural judgements in the perception of computer-generated music. In *Second International Conference of Asia Pacific Society for the Cognitive Science of Music*. Seoul, Korea, pp. 155-160.

Bailes, F., & Dean, R. T. (2007a). Facilitation and coherence between the dynamic and retrospective perception of segmentation in computer-generated music. *Empirical Musicology Review*, Vol. 2, No. 3, pp. 74-80.

Bailes, F., & Dean, R. T. (2007b). Listener detection of segmentation in computer-generated sound: An experimental study. In press, *Journal of New Music Research*.

Brattico, E., & Sassanelli, F. (2000). Perception and musical preference in Wishart's work. *Journal of New Music Research*, Vol. 29, No. 2, pp. 107-119.

Cross, I., & Rohrmeier, M. (2007). Comments on 'Facilitation and coherence between the dynamic and retrospective perception of segmentation in computer-generated music', by Freya Bailes and Roger T. Dean. *Empirical Musicology Review*, Vol. 2, No. 4, pp. 146-148.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, Vol. 132, pp. 87-110.

Van Nort, D. (2007). *Texture perception: Signal models and compositional approaches*. Paper presented at the conference of the Society for Music Perception and Cognition, July 30-August 3, Montréal, Canada.