

**Comments on “Facilitation and Coherence Between  
the Dynamic and Retrospective Perception of  
Segmentation in Computer-Generated Music,”  
by Freya Bailes and Roger T. Dean**

IAN CROSS

*Centre for Music & Science, Faculty of Music, University of Cambridge*

MARTIN ROHRMEIER

*Centre for Music & Science, Faculty of Music, University of Cambridge*

**ABSTRACT:** Although the study by Bailes & Dean (2007) addresses an under-researched area of auditory and musical perception, it raises questions concerning stimuli, methodology, and the study's relation to previous research, that are outlined in this commentary.

Submitted 2007 August 20; accepted 2007 August 20.

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

BAILES and Dean (2007) have identified a significantly neglected area of research, the perception of timbre and texture and its impact on perceived segmentation (particularly aspects that might be relevant for contemporary electroacoustic composition), and have taken initial steps towards its exploration. They created complex sound stimuli, presented these in pairs as continuous signals, and explored the effects of stimulus make-up and duration on listeners' judgments of whether or not the signals were continuous or segmented and, if segmented, the point at which segmentation was perceived to occur. They found that listeners were generally accurate in perceiving segmentation, and that the points of segmentation were generally more accurately located when the duration of the first stimulus of a pair exceeded that of the second. They interpret the results as indicating that stability of representation of the initial sound of a pair in memory facilitated segmentation judgments, pointing up (p. 79) “the importance of considering the immediate sound context in the perception of structure”.

After a flurry of research in the 1970s and 1980s (see, e.g., Grey & Moorer, 1977; for a review see Risset & Wessel, 1999), research on timbre appears to have receded in prominence in the published literature (with a few significant exceptions such as the work of Stephen McAdams: see, e.g., Caclin, McAdams, Smith & Winsberg, 2005). However, the issue of the relationship between perception and contemporary composition—and indeed, that of the relevance to contemporary compositional practice of an understanding of the perceptual processes involved in listening—remains live and likely to be of considerable importance to anyone seeking to deal with the array of possibilities opened up by new technologies and ways of conceptualising sound in and as music. Nevertheless, the results of the experiment reported here seem unlikely significantly to advance an understanding of timbre and texture perception.

We feel that there are three main aspects of the study that require comment. The first, and perhaps the most fundamental, issue is the way in which the study locates itself in the field, particularly in terms of its characterisation of previous and potentially relevant research. The second relates to the ways in which

stimuli are themselves characterised and reported. The third concerns methodological considerations and the form in which the experiment was conducted.

In regard to the portrayal of the research area in the present paper, there is much more research on the perception of timbre and texture than is evident from the context provided here; although the literature is widely dispersed, in addition to the papers cited above a good overview of issues and findings germane to current research is provided in Handel (1995). One issue that arises from the lack of reference to prior timbre research here is that of the conception of timbre that the authors adopt; this impacts on the scope of the inferences that may be drawn from the results of their experiment. Current understandings of the nature of perceived timbre suggest that there are three possible ways of conceptualising it. The experience of timbre can be thought of as deriving from *innate* understandings of the ways in which the acoustical attributes of a signal necessarily derive from the properties and dynamics of the mechanical systems producing the signal (see, e.g., Gaver, 1993). A second possibility is that perceived timbre is based on *learned* understandings of the relationships between acoustic features and mechanical properties. The third view would take as its starting position the abstract characteristics of the sound structures (as in 'acousmatic' listening) and suggest that the experience of timbre will correlate in some ways with these structures. What is currently known would suggest that the first two possibilities should be favoured, conceiving of timbre as perceived in terms of feature of real-world actions and events (whether innate or learned); however, the issue of the possibility of 'acousmatic' listening still remains to be comprehensively addressed in experiment (though see, e.g., Windsor, 1995; Brattico & Sassanelli, 2000). The present study, in referring to the effects of auditory looming and to Huron's RAMP archetype, appears implicitly to reject the acousmatic approach, but it would have been helpful for the authors' position to be clarified.

In terms of methodology, the experiment reported here uses stimuli created by specialist electroacoustic software (such as Max/MSP), but does not describe these in ways that enable their acoustical correlates to be identified and understood (it is noted (p. 77) that "Many of the individual sound segments are time-variant, and consequently the inter-segment relationships are not easy to quantify."). This limits the scope for interpretation of the results and replicability of the experiment. Moreover, the number of stimuli employed was low (36), as was the number of experimental participants (14); experiments exploring perceptual thresholds and limits will usually either involve large numbers of stimuli, or of participants, in order to ensure statistical interpretability and validity of results. A typical experiment exploring the ability of listeners to segment pairs of stimuli that are presented as continuous events (as in the present experiment) would tend to present multiple instances of each stimulus. It would also tend to vary exhaustively the relationship between the stimuli of each pair by small steps in respect of parameters that are acoustically specifiable. The purpose of this would be to explore detection thresholds by means of an adaptive presentation strategy (such as that proposed in Levitt, 1971), expressing those thresholds in terms of Signal Detection Theory (detection thresholds in the current study would be the values of stimuli for which segmentation (i.e., change) is consistently perceived to occur, which would require multiple presentations of each stimulus).

However, it would certainly be the case that recasting the present study in these terms would necessarily introduce some logistical difficulties. Multiple presentations of stimuli would greatly extend the duration of each test, but would provide more robust data that would have been much more securely interpretable in statistical terms. Exhaustive varying of parameter values in an adaptive task would again greatly lengthen the practical duration of each experimental session, raising the danger of fatigue effects and making participant recruitment difficult (hence a psychoacoustical experiment will often make use of a very small number of participants—not atypically, as many as there are authors of the paper...). Nevertheless, it would seem worthwhile to make the effort; the issues that are being explored in the programme of research outlined here are of considerable interest. A study or studies that focused on these issues yet took more account of classical psychophysical methods would have the potential to yield insightful and informative results that would be crucial for our understanding of timbre and texture perception and valuable for the contemporary practice of musical composition and performance.

## REFERENCES

- Bailes, F., & Dean, R.T. (2007) 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.
- Brattico, E., & Sassanelli, F. (2000). Perception and musical preferences in Wishart's work. *Journal of New Music Research*, Vol. 29, No. 2, pp. 107 - 119.
- Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America*, Vol. 118, No. 1, pp. 471-482.
- Gaver, W. W. (1993). How do we hear in the world? Explorations in ecological acoustics *Ecological Psychology*, Vol. 5, No. 4, pp. 285-313.
- Grey, J. M., & Moorer, J. A. (1977). Perceptual evaluations of synthesized musical instrument tones. *Journal of the Acoustical Society of America*, Vol. 62, No. 2, pp. 454-462.
- Handel, S. (1995). Timbre perception and object identification. In B. C. J. Moore (Ed.), *Hearing*. London: Academic Press, pp. 425-461.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, Vol. 49, No. 28, pp. 467-477.
- Risset, J. C., & Wessell, D. L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed.). London: Academic Press, pp. 113-169.
- Windsor, W. L. (1995). *A Perceptual Approach to the Description and Analysis of Acousmatic Music*. Ph.D dissertation, City University, London [available at: <http://www.personal.leeds.ac.uk/~muswlv/pubs/files/wlwthesis.doc>]