

# AN ANALYTICAL METHODOLOGY FOR ACOUSMATIC MUSIC

David Hirst

Teaching, Learning and Research Support Dept, University of Melbourne

email: d.hirst@unimelb.edu.au

## ABSTRACT

This paper presents a procedure for the analysis of acousmatic music which was derived from the synthesis of top-down (knowledge driven) and bottom-up (data-driven) cognitive psychological views. The procedure is also a synthesis of research on primitive auditory scene analysis, combined with the research on acoustic, semantic, and syntactic factors in the perception of everyday environmental sounds. The procedure can be summarized as consisting of a number of steps: Segregation of sonic objects; Horizontal integration and/or segregation; Vertical integration and/or segregation; Assimilation and meaning.

## 1. INTRODUCTION

Not only do we recognize sounds, but we ascribe meanings to sounds and meanings to the relationships between sounds and other cognate phenomena. Music is a meaningful and an emotional experience. Acousmatic music uses recordings of everyday sounds, recordings of instruments and synthesized sounds. They are combined in their raw forms or processed then mixed together and recorded to a fixed medium (usually CD or tape). Acousmatic music may make use of traditional musical relationships, but there can also be unique abstract relationships between the sonic attributes of sounds and the perceiver of those sonic attributes that we don't find in traditional instrumental music. What is this "syntax" of acousmatic music and how does it interact with the semantic references afforded by some of the sonic material within acousmatic musical works?

We have previously reported on a bottom-up type of approach to the analysis of acousmatic music that adapted Bigand's model of Event Structure Processing of tonal music (Hirst, 2003) [5], and discussed the interpretation of some of Smalley's concepts (Hirst, 2002) [4]. In this paper we will examine some of the top-down processes that may operate in the perception of Western tonal music and also consider top-down processes in the recognition of environmental sounds. Through combining these two bodies of previous work, where there has been some research activity, we may shed some light on the perception of acousmatic music, where there has been only a meagre amount of research. The combination has resulted in the definition of a methodology for the analysis of acousmatic works.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2004 Universitat Pompeu Fabra.

## 2. DEFINING THE ANALYTICAL METHODOLOGY

The following procedure for the analysis of acousmatic music was derived from the synthesis of top-down (knowledge driven) and bottom-up (data-driven) views since it has been found impossible to divorce one approach from the other.

The procedure can be summarized as consisting of a number of steps:

### *Segregation of sonic objects*

1. Identify the sonic objects (or events).
2. Establish the factors responsible for identification (acoustic, semantic, syntactic, and ecological) and their relative weightings (if possible).

### *Horizontal integration and/or segregation*

3. Identify streams (sequences or chains) which consist of sonic objects linked together and function as a unit that we could call a "pattern". The role of "trajectories" should also be considered. (The term "gesture" has sometimes been used in this context.)
4. Determine the causal linkages between the sonic objects within the chain-type pattern.
5. Determine the relationships between "pattern objects" – if this level of syntax exists. This amounts to the investigation of higher-order relationships within a "hierarchy".
6. Consider local organization in time – pulse, beat, accent, rhythm, meter.
7. Consider the horizontal integration of pitch, including emergent properties relating to timbre (vertical overlap).

### *Vertical integration and/or segregation*

8. Consider vertical integration as a cause of timbre creation and variance:
  - a. Timbre as a cause of integration and/or segregation.
  - b. The dimensional approach to timbre, including emergent properties relating to pitch (horizontal overlap).
  - c. Texture resulting from contrasting timbres.
9. Also consider vertical integration or segregation in terms of the potential for psychoacoustic dissonance and musical dissonance (and consonance).

### *Assimilation and meaning*

10. Consider the nature and type of discourse on the source-cause dominant to typological-

relational dominant continuum, and the way it varies over time.

11. Consider implication-realization, and arousal and meaning on a moment-to-moment basis throughout the work.
12. Consider global organization in time – identify formal structures, like sectional or continuous organization, and the nature of the relationships between sections, i.e. hierarchical relationships.

### 3. APPLYING THE ANALYTICAL METHODOLOGY

Having outlined a methodology, and considered some preliminary issues, let us examine each of the steps in the above procedure in more detail.

#### 3.1. Segregation of sonic objects

The first stages in our methodology are the identification the sonic objects (or events), and establishing the factors responsible for identification.

Building on the work of Bregman, McAdams has postulated several stages in the recognition of sound sources and events (McAdams, 1993) [8]. The first two stages are sensory transduction and auditory grouping of frequencies – both simultaneous and sequential. The third stage involves the analysis of auditory properties and features.

There may be many factors responsible for categorical matching so the next step in our methodology seeks to explore the varied factors and the possible contribution made by each factor in sound event recognition.

##### 3.1.1. Factors in the identification of environmental sounds

###### *Acoustic factors*

Gygi (2001) presents results of similarity studies conducted at the Hearing and Communication Laboratory at Indiana University. He points out that for “meaningful” sounds, there needs to be a consideration of their “psychological space”. [3]

One experiment attempted to uncover the structure of such a psychological space by using multidimensional scaling (MDS) procedures.

Gygi conducted the experiment using 50 environmental sounds. Recordings of 100 instances of the environmental sounds were played to listeners who were asked to rate the similarity of each pair on a seven point scale from least similar to most similar.

A two-dimensional MDS solution was derived. Gygi searched for a meaningful interpretation of the two dimensions. He measured over twenty acoustic variables covering different aspects of spectral distribution. The highest correlations for Dimension 1 were spectral spread, and confidence in the pitch of the signal. Correlations with the second dimension were not so

clear cut. The most significant correlations were to do with measures of rhythmicity and spectral mean.

Gygi’s identification studies gives us a basis for isolating acoustic features of environmental sounds that we may want to single out for analysis.

###### *Syntactic and semantic factors*

In an article by Howard and Ballas (1980), they documented their experiments which tested syntactic (temporal structure) and semantic factors (knowledge of source events) in the classification of non-speech transient patterns [6]. What is important about this series of experiments is that they are concerned with *patterns* of sounds (i.e. sequences of sounds) – both pure tones and environmental sounds.

Howard and Ballas found that a grammatical group of listeners performed significantly better than did a non-grammatical group for pure tones and for semantically sensible sequences. Listeners in the grammatical group had actually learned something about the syntactic rules (the grammar) used to generate the target patterns.

#### 3.2. Horizontal integration and/or segregation

In this next phase of our analytical procedure we must identify sequences (streams or chains). Such a sequence would consist of sonic objects linked together and it would function as a single unit that we could call a “pattern”. The role of “trajectories” should also be considered. (The term “gesture” has sometimes been used in this context.)

Next we need to determine the linkages between the sonic objects within each chain-type pattern. Which of the relevant factors are operating, and what are the relative weightings of those factors?

Then we shall determine the relationships between “pattern objects” – if this level of syntax exists. This amounts to the investigation of higher-order relationships within a “hierarchy”.

Finally we must consider local organization in time – pulse, beat, accent, rhythm, and meter.

This phase is a lot like a melodic analysis in tonal music. Transformations in pitch or frequency define a melodic form (or gesture). However, we shall treat pitch as a quality that is emergent from the frequency spectrum of sonic objects within acousmatic music.

A sequence with large frequency transitions in a short space of time will not remain perceptually coherent. Alternations of high and low frequencies will cause segregation between the two different registers and the perception of separate streams will result.

Changes in timbre can effect the integration of a horizontal sequence, for example: Repeated and/or rapid changes in timbre can fragment a sequence; Less rapid shifts in timbre can be used to delineate larger horizontal units or “phrases”.

A distinction can be made between form-defining sequential events and subordinate ornamental sound events. Ornamental events group with the form-bearing event they are subordinate to. Gestalt-like principles of

primitive scene analysis seem to constrain these emergent events. The dependent event must be very close to the anchor event in frequency and time, otherwise the two will not group to form a larger event.

The choice of which event is the anchor and which is the dependent one depends on factors such as duration, intensity, and rhythm. Dependent events tend to “resolve” to stable anchor events.

### 3.2.1. Trajectories

Is a regular sequence of sounds (where the sequence is predictable) easier to recognize than an irregular one?

As we listen repeatedly to an auditory pattern, we learn the regularities in it. We can then anticipate segments of the pattern before they occur and integrate the sequence into a coherent mental representation. This preparation of our attention through repeated listening should assist the formation of sequential streams.

There is evidence both for and against abilities in the formation of streams via trajectories. This concept of *trajectory* is important, but its application is far from straight forward. (See Bregman, 1999: 670) [1]

## 3.3. Vertical integration and/or segregation

In this next stage we consider vertical integration as a cause of timbre creation and variance. We also consider vertical integration or segregation in terms of the potential for psychoacoustic dissonance and musical dissonance.

### 3.3.1. Timbre and texture

Bregman notes that timbre plays a role in the sequential organization of music. Timbre is a complex phenomenon and no one perspective can encapsulate its complexity. Bregman tackles timbre from several points of view.

#### *Timbre as a cause of segregation and integration*

As we saw above, large and rapid changes in timbre can cause the formation of separate, parallel streams as happens with a “compound melody”.

Conversely, timbre can be used as sequential glue for musical phrases, and to delineate separate units or sections. Such delineation exploits scene analysis principles from nature where a sudden change in timbre usually implies a new event has begun. In contrast, a continuous change implies that a single event is changing in some way – through some form of incremental transformation.

#### *The dimensional approach to timbre*

Bregman points out that there is a use of timbre that is stronger than the accentuation and reinforcement of melodic forms.

Bregman highlights several approaches to the study of the dimensionality of timbre. Lerdahl (1987) has attempted some tests of the organization of music

according to certain types of dimensionality [7]. Other studies have explored acoustic dimensions such as the formant frequencies of vowel sounds. (See Slawson, 1985) [9].

#### *Timbre as the result of fusion, texture as a result of contrasting timbres*

Knowledge of fusion and segregation principles can be used to assist the organization of musical texture, for example a number of musical sounds can be fused to create a global timbre, or a polyphonic texture can be created where two or more distinct sequential streams can be heard.

Tonal music uses scene analysis principles in a number of ways that are relevant to musical texture. Perhaps these principles could be adapted for the analysis of texture in acousmatic music.

An examination of the role of primitive scene analysis in counterpoint may be of some benefit too. In polyphonic music, the parts must not be totally segregated or totally integrated. Segregation between parts is improved by strong sequential organization within each part. Strong sequential organization is achieved through principles such as: using small pitch changes to favour sequential integration; avoiding “common fate” by prohibiting synchronous onsets and offset (different rhythms); avoiding parallel changes in pitch between two parts, i.e. encourage contrary motion or oblique motion; avoiding harmonic relations between simultaneous events in different parts, e.g. whole number ratios like 2:1, 3:2, etc.

#### *Dissonance*

Bregman distinguishes two types of dissonance: psychoacoustic dissonance and musical dissonance. Psychoacoustic dissonance is the sense of roughness caused when partials combine to produce a large number of beats at an unrelated rate [2]. Simple ratios like 3:2 (ca. seven semitones) will sound smooth, complex ratios like 45:32 (ca. six semitones) will seem rough.

Musical consonance is defined by Bregman as a cognitive experience. Stable and unstable combinations of sounds are defined by a musical style. Unstable combinations are points of tension, whereas stable ones are points of rest. Bregman notes that unstable combinations of tones often happen to be psychoacoustically dissonant in Western music.

Composers control such dissonant tones by using techniques such as: avoiding simultaneous start and stop times; capturing tones into separate streams by preceding each with tones close in pitch; capturing each into smooth and different trajectories; capturing each into its own repetitive sequence. An increase in dissonance can be created by violating the above principles.

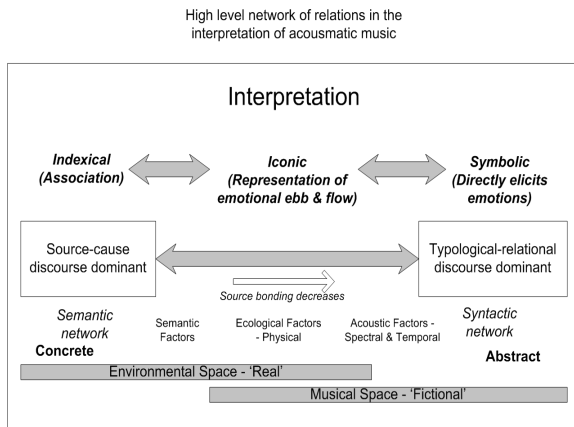
## 3.4. Assimilation and meaning

In this final stage we consider global organization and formal structures. This is also the stage in which we

discuss emotion, arousal and meaning, implication-realization, and the nature and type of discourses operating within the work. Space does not allow us to consider these concepts within this paper, but the model displayed in Figure 1 draws upon the work of Dowling and Harwood (1986) [2].

#### 4. CONCLUSION

The above approach is a reductionist one – tearing the musical materials apart. We must now put the work back together again in order to get a more holistic view.



**Fig. 1:** Network of relations in the interpretation of acousmatic music.

Figure 1 is a pictorial representation of the concepts introduced in this paper and their relationships. Of fundamental importance is the differentiation between “Environmental Space” associated with the experience of “real-world” sounds and “Musical Space” associated with the “fictional sound world” created by the musical work. Note that there is scope for overlap between these worlds so that the composer can play with ambiguities and chimerical properties resulting from their combination.

In assessing the dominant forces at play within a work, a continuum has been constructed according to the type of discourse that may be dominant at any given time. Source-cause discourse (after Smalley, 1994) is dominant when sounds from the natural world are identified and related somehow within a semantic network. [10] These sounds will have strong source-bonding and the syntax of their relationships will be dominated, although not exclusively, by semantic and ecological factors.

As source-bonding decreases, the acoustic factors begin to dominate and the work adopts a typological-relational discourse based upon more abstract relationships within a syntactic network governed by implications and realizations. The continuum extends from the purely concrete reproduction of the

environment (real space) to the totally abstract world of a fictional sound space with fictional sources (musical space).

References extend from the indexical to the symbolic, from simple association at one extreme to directly eliciting emotions at the other extreme, with a representation of emotional ebb and flow in between. The listener’s experience along these various continua can vary from moment to moment within a given work, and this becomes one more dynamic aspect within the discourse.

What remains to be completed is to test this model of analysis on representative repertoire works.

#### 5. REFERENCES

- [1] Bregman, A. S. (1999). *Auditory Scene Analysis: The Perceptual Organization of Sound*. (Second MIT Press Paperback edition) Cambridge, MA: MIT Press.
- [2] Dowling, W.J. and Harwood, D.L. (1986). *Music Cognition*. Orlando: Academic Press.
- [3] Gygi, B. (2001). *Factors in the Identification of Environmental Sounds*. PhD Dissertation: Indiana University
- [4] Hirst, D. (2002). Developing Analysis Criteria Based on Denis Smalley's Timbre Theories. *Proceedings of the Australasian Computer Music Conference 2002*, pp 43-52. Australasian Computer Music Assoc. Melbourne.
- [5] Hirst, D. (2003). Developing a cognitive framework for the interpretation of acousmatic music. *Converging Technologies: Proceedings of the Australasian Computer Music Conference 2003*, pp 43-57. Australasian Computer Music Assoc. Melbourne.
- [6] Howard, J. H. & Ballas, J. A (1980). Syntactic and semantic factors in the classification of nonspeech transient patterns. *Perception & Psychophysics* 28(5), 431-439.
- [7] Lerdahl, F. (1987). Timbral hierarchies. *Contemporary Music Review* 2: 135-60.
- [8] McAdams, S. (1993). Recognition of sound sources and events. *Thinking in Sound: The Cognitive Psychology of Human Audition*. (ed. S. McAdams and E. Bigand). Oxford University Press, pp. 146-98.
- [9] Slawson, W. (1985). *Sound Color*. University of California Press, Berkeley, CA.
- [10] Smalley, D. (1994). Defining Timbre - Refining Timbre. *Contemporary Music Review* Vol. 10 Part 2, Harwood Academic Publishers, Switzerland, pp. 35-48.