

Auditory Scene Analysis and Electroacoustics

Rosemary Mountain
Music Department, Concordia University
7141 Sherbrooke St. West, RF 322
Montreal, Canada H4B 1R6
mountain@alcor.concordia.ca

ASA and EA

1. Tracing the author's encounters with and incorporation of Auditory Scene Analysis

I first discovered the field of auditory perception in the mid-1980's, as I began playing with the idea of using polyrhythms as a basis for large-scale musical structure. Issues of auditory and temporal perception seemed particularly relevant for examining the difference between small- and large-scale structures, and I quickly began to rely on perception & cognition research for exploring many issues relating to auditive analysis and, in general terms, the influences on our listening experience. Eventually, I also began consulting that research for guidance on composition, as it seemed useful to understand potential tendencies of and challenges to listening that could arise from particular sonic configurations. In particular, the field known as Auditory Scene Analysis (or ASA) developed by Bregman and hundreds of students and colleagues over the last forty years, has become a mandatory reference for most of my teaching, as I find elements of its findings relevant for courses in music analysis, composition, and electroacoustics. Usually I include some reference to cognate areas as well, such as information theory, which involves understanding of such useful and closely-related concepts and terms as signal:noise ratio, bits of information, redundancy, and short-term memory limits. Information theory in particular has fed into electroacoustics for many years, thanks to the expertise of Meyer-Eppler, who helped establish the Cologne school and influenced Stockhausen directly; Ligeti was also very aware of perceptual issues and mentions their impact on his compositional ideas (see for example Ligeti 1983). One aspect of Auditory Scene Analysis that makes it particularly compatible to electroacoustic studies is that, as it emerges from a study of sound, it uses familiar terms such as frequency, spectrum, amplitude, etc. whereas its relevance to the acoustic musician may need more 'translation'. In addition, considerable overlap exists between Auditory Scene Analysis and Gestalt psychology, as applied to visual arts and subsequently to music. Leonard Meyer served to introduce many of my generation in music to those concepts (see for example *Emotion and Meaning in Music*, 1957), which include figure-ground relationships and the 'laws' of 'good continuation', 'proximity', and 'similarity'. A major difference between the ASA approach and that of Gestalt psychology is that while Gestalt psychology describes the results of our grouping tendencies, ASA attempts to investigate what it is in our auditory-cognitive apparatus that causes us to perceive sounds in ways which reveal those groupings.

The essence of ASA research is in fact an investigation of groupings, or the way in which we "parse" sounds, to borrow from linguistics. Parsing, in a general sense, refers to our ability to group phonemes into words, and words into phrases, and phrases into sentences, to render intelligible a stream of sounds. Likewise, ASA allows us to understand how we are likely to group different elements of a sonic array into meaningful structures and patterns. It is this aspect which led me to incorporate ASA into a new tool to help analyze complex works which appear to have several different layers (as explained in Mountain 1998 and 2001). I taught this approach in the context of advanced analysis courses while at the University of Aveiro (Portugal); the class worked in small groups to analyze several pages each of a complex work. Over the course of three years, different students completed preliminary analyses of Stravinsky's *Rite of Spring*, Messiaen's *Turangalila*, Ives *Fourth Symphony*, and Varèse's *Amérique*. The students identified the different component strata aurally, then made an "inventory" of the components of each one (pitch collection, range, dynamics, durations, etc). They examined factors which seemed to promote cohesion within each stratum, and those which led to the perception of segregation (see Tables I and II).

When I arrived at the Department of Music at Concordia University (Montreal, Canada), I naturally introduced basic concepts of ASA into the curriculum of the introductory electroacoustics course. However, I did not think of applying it to analytical discussions, as my previous work with ASA for analysis had been focussed on identifying potentially audible structures in written scores. In addition, I had a sense that the electroacoustic composer's layering of material into different tracks already led to enough differentiation that the challenge of identification was less marked (more on this below). It was therefore interesting to find ASA quickly becoming part of the electroacoustic curriculum at Concordia; I attributed this mainly to the obvious usefulness of understanding issues of human perception relating to sound, rather than to the analytical potential of ASA. However, recently I have been reflecting on the potential of integrating ASA with research such as 'les Unités Sémiotiques Temporelles' (UST) developed by the group MIM (Laboratoire Musique et Informatique de Marseille), and the set of terms, musical examples, and classifications delineated by Marcellehènes (Univ. de Montréal), to formulate potent analysis tools. Through the use of flexible notation tools such as the *Acousmographie* and Couprie's *iAnalyse*, these characteristics can be more clearly identified.

2. Bibliographic review

The most comprehensive overview of the field of ASA is found in the 790-page book *Auditory Scene Analysis* written by Al Bregman, pioneer of the area (1990); this was greatly enhanced by the subsequent publication of the CD (with Pierre Ahad - 1996) which gives audio examples of several experiments and allows the listener to hear first-hand some of the phenomena, as well as to appreciate the types of experiments conducted, in terms of sonic examples and presentation. However, for those not used to psychological methodology, and/or not overly concerned with the details of the experiments required to test various manifestations of perceptual phenomena, Bregman's chapter (1993) in the Bigand/McAdams book *Thinking in Sound* is a good starting point, and his newly-revised webpage (2008) is also invaluable as it provides mp3 versions of the experiments from the CD along with annotations and an overview. My own introduction to the field was not the book (which was not yet published) but rather dozens of articles structured according to standard psychological reporting, with the vast majority of their content devoted to details of the experiment, a brief summary giving results and only sometimes hypothesizing some broader conclusions; I was therefore delighted to receive a summary in the form of a two-page handout distributed by Bregman at an International Musicology Society conference in 1997 (Bregman 1997); some of this is adapted in Table I. As Bregman's research has been concerned with the way in which sounds of all types are sorted by the mind, the relevance for music is not always obvious. Therefore, I was particularly pleased to find two articles by Stephen McAdams (1982, 1987) which help clarify the links between ASA research and music.

1. Sequential grouping:	2. Simultaneous grouping (overlap in time):
<u>Factors that increase segregation:</u> <ul style="list-style-type: none">a) frequency separation,b) separation in timec) differences in timbre,d) differences in spatial origin,e) differences in loudness (weak effect),f) abruptness of the transition from one sound to the next.	<u>Factors that increase segregation:</u> <ul style="list-style-type: none">a) different harmonic relations among subsets of partials (common fundamental frequency);b) different places of origin in space;c) frequency separation;d) independent changes in loudness (synchronized dynamics);e) asynchrony of attack (and to a lesser degree, shutoff) of frequency components;f) asynchronous changes in frequency (maybe).

Table I. Acoustic clues for the correct grouping of parts (from Bregman 1997)

3. Concepts of ASA and their relationship with music

Auditory scene analysis is the study of how we try to segregate the component parts of a sonic environment 'back' into their individual sources. A visual analogy is given in Figure 1(a) and (b): a rather undifferentiated set of dots can be read as a much more intricate construction when the dots are given different shades of grey, and the mind assumes connections between them, forming lines and patterns, even when those patterns are not portraying anything familiar. In the case of a non-musical context, a typical situation of ASA could involve distinguishing two human voices, a car, a dog barking, and an air conditioner; an easier feat to accomplish than to explain. However, in musical contexts, it can

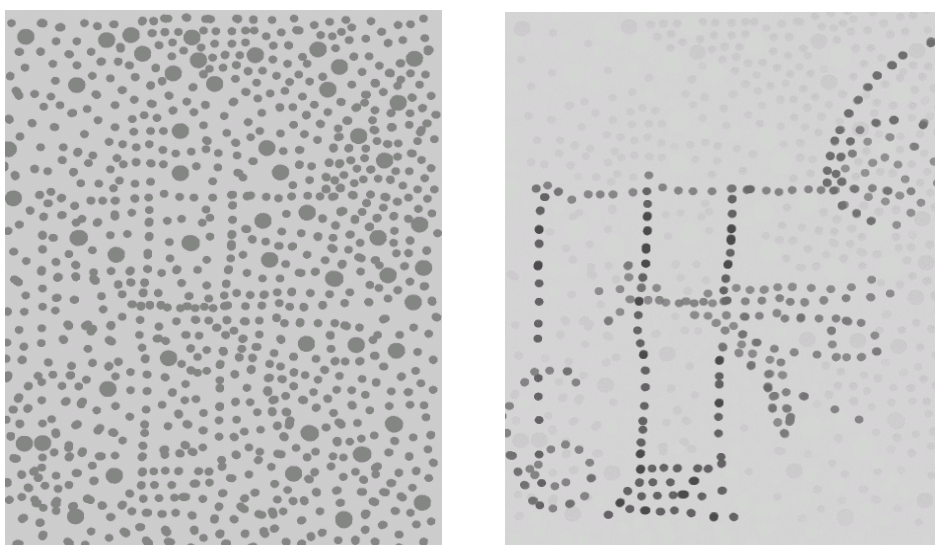


Figure 1. Two versions of the same texture: (a) largely undifferentiated, and (b) with changes in grey scale.

be more complex and more ambiguous. For example, in an orchestral texture, we may be able to distinguish the oboe

from the trombone - but perhaps that is irrelevant to the musical structure, if they are meant to be jointly playing a composite harmonic line. And in many contemporary works, the composer deliberately creates situations where several sounds are meant to be heard as a fused entity and then gradually metamorphose into separate strands - or vice versa; see Figure 2. (Early Ligeti works abound with such passages.) The question is, therefore: which factors will influence us to hear two or more sounds as belonging to a single 'voice' or texture - or 'fusion' - and which will influence us to hear them as distinct from each other? Bregman's guides (Table I) distinguish between simultaneous and sequential differentiation, or what in music might be referred to as superposition and juxtaposition. However, although in a laboratory the differences between sequential and simultaneous are clear, in musical contexts they can seem more blurred, as in Bach's use of implied polyphony in the solo violin and solo cello suites; quick alternation between notes is strictly speaking sequential, but gives the impression of two simultaneous lines the temporal focus, or window, is just a bit broader. Also, the mind is constantly trying to identify sounds as being from one or more sources, so corrections may need to be made as new information emerges: Figure 2 illustrates a passage where the first part (2a) might appear as a single texture, but subsequent information (2b) indicates that it is quite different. Here again, the time scale is important, because if the Xs refer to (groups of) notes of rich timbre, and 2(a) is a few seconds or so in length, we may consider it in retrospect a 'wrong' reading, whereas if it is a minute in length, we may consider it a metamorphosis; if the Xs refer to simple frequencies, we may consider 2(b) an example of spectromorphology!

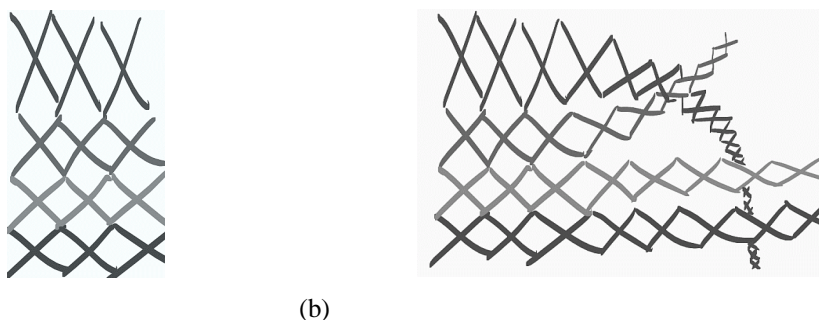


Fig. 2. A passage may be understood as a single entity (a) but later reinterpreted as the initiation of several layers (b).

My experience is that often a passage will contain factors that promote segregation of two lines, and other factors that promote fusion - therefore, it becomes a question of calculating the respective weights of these opposing influences.

In Figure 3, we can see a representation of a fairly typical situation in music, if we read it like a score with frequency as the vertical axis and time as the horizontal one. In terms of auditory scene analysis, we recognize that the top and middle figures would tend to fuse due to onset synchrony and contour, but registral distance would tend to segregate them (though not if they are harmonically related by an octave, for example). On the other hand, the lower two lines will have some tendency to fuse due to register, and somewhat due to contour (but in different timescale), but their onset asynchrony will lead to a tendency to segregate.

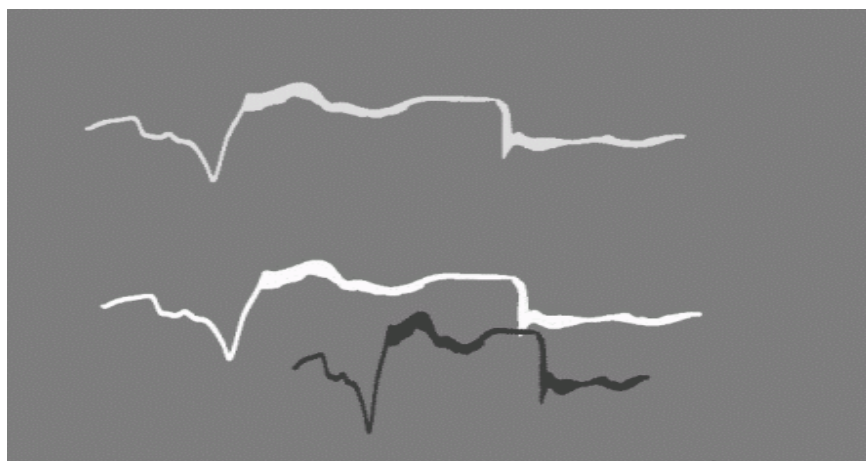


Fig. 3. Conflicting factors of fusion and segregation

Obviously, such calculations are not (at the moment anyway) able to be very rigorous or scientific; instead, one must rely on a musical sense, often involving analysis and a sense of the composer's aesthetic. In some cases, the listener can choose to hear a passage in one or another way, and performance can often become a major influence as the factors on one side are reinforced.

The easiest way to grasp the tendencies of fusion and segregation is to consider that our perceptual processing has its roots in very ancient adaptation to the environment: optimized for distinguishing voices of friends and enemies (of

human and animal kind) from other sounds such as fire, water, and wind. The human voice is rich in spectra but the behavior of the partials is united by the breath (amplitude variations), the mouth cavity, and the spatial location even as the vowels and consonants change. Likewise, as McAdams (1982) points out, when several instruments apply a common behavior such as dynamics and rubato to a succession of notes, they will appear to emanate from a common source. In electroacoustic music, this effect of a single voice is often achieved (sometimes inadvertently) through the treatment of separate 'layers' by placing them in different tracks and applying common processing. The brain is evidently very sensitive to slight differences of treatment implying different sources, and conversely will tend to assume a common source even when disparate elements are subjected to common treatment such as panning, amplitude, echo, etc.

4. Clarification of terminology

	FUSED	DISTINCT
ONSET	simultaneous	varied
SPATIAL LOCATION	near	far
REGISTER	similar	far apart
SUPER-PULSE	identical	different
CONTOUR	similar	contrasting
TIMBRE	similar	contrasting
PULSE	identical	different
SUB-PULSE	simply related (1:1, 2:1, 3:1)	dissonant

Table II. Factors affecting perception of two (or more) sets of sonic elements as being grouped together ('fused') or segregated ('distinct').

Table II shows my own chart of factors arranged in a slightly different manner to Bregman's, and with more music-specific terms. The order of the features relates very roughly to a typical order of influence. 'Onset' means the beginning of two or more elements: notes, partials, phrases. By 'onset asynchrony' we refer to the lack of simultaneous attack between two or more elements. This may occur in a large-scale situation, such as the staggered entries of a fugue, or at micro level, such as the slightly-delayed entries of all accompanying voices in a piano texture which allow the listener to hear the melody as the principal voice, even when in an "inner voice" (and therefore not attracting our attention by virtue of a privileged top voice of a texture). The term 'super-pulse' is one I coined while working on my dissertation (Mountain 1993), to refer to the periodicity level of around 1.5 - 6 seconds, that is, around the usual range of a measure (or downbeat-to-downbeat cycle) in metric organization; this is an area where we seem to be quite sensitive to deviation, with different reactions than to deviations in the sub-pulse level (which can be sensed more as a communication of expression). Contour appears very influential, but again the context and the listener's experience will influence the degree of similarity perceived. It seems logical that timbre would have an important role in differentiating between layers, but on reflection, Western music at least seems to discourage this role: we are expected to treat a theme as "the same" regardless of its orchestration.

I am quite convinced that a good grounding in Auditory Scene Analysis, and in auditory and temporal perception in glued, can be invaluable for all musicians, whether performers, composers, teachers, or theorists, and whether in acoustic or electroacoustic music. I doubt that we will develop a "science" as there are many ambiguities, but the issues studied are very relevant for explaining tendencies of listening. In addition, I believe that musicians could contribute some valuable insights into the field, as we deal with illusions and constructions in ways which stretch the boundaries of human perception, and it is good for psychologists to remember that! The current format gives me little room for full explanations of my findings to date, but I hope to publish a much more detailed report in the near future.

Bibliography:

- BREGMAN, Albert, Auditory Scene Analysis, Cambridge, MA, MIT Press, 1990.
 BREGMAN, Albert, "Auditory Scene Analysis: Hearing in Complex Environments", in: Stephen McAdams and Emmanuel Bigand (eds.), Thinking in Sound: The Cognitive Psychology of Human Audition, Oxford, Oxford University Press, 1993, 10-36.
 BREGMAN, Albert, "Auditory Scene Analysis: Hearing in Complex Environments", in: Stephen McAdams and Emmanuel Bigand (eds.), Thinking in Sound: The Cognitive Psychology of Human Audition, Oxford, Oxford University Press, 1993, 10-36.
 BREGMAN, Albert, and Pierre AHAD, Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound, Cambridge, Mass., MIT Press, 1993 (audio CD).
 BREGMAN, Albert, "When Will We Hear Separate Events in a Sequence of Sounds?" handout accompanying his talk at the International Musicological Society, London, 1997.
 BREGMAN, Albert, "Al Bregman's Website", <<http://webpages.mcgill.ca/staff/Group2/abregm1/web/>> 2008.

- LIGETI, György, *Ligeti in Conversation with Péter Várnai, Josef Häusler, Claude Samuel and himself*, London, Eulenberg, 1983.
- McADAMS, Stephen, "Spectral fusion and the creation of auditory images" in: Manfred Clynes (ed.) *Music, Mind, and Brain*, 1982, 279-298.
- McADAMS, Stephen "Music: A science of the mind?" 1-61, in: *Contemporary Music Review* vol. 2, no. 1, 1987.
- MEYER, Leonard, *Emotion and Meaning in Music*, Chicago, University of Chicago Press, 1956.
- MOUNTAIN, Rosemary, "Sorting out the Strata: Auditory Scene Analysis Applied" presented at the Dept. of Music, Univ. of Ottawa and to Dr. Al Bregman's graduate seminar, Psychology Dept, McGill University, Montréal, 1998.
- MOUNTAIN, Rosemary, "Sorting out the Strata - Revisited" – presented at the Music Cognition Seminar of the Brain and Cognition Department, University of Rochester; the Psychology Department, Cornell University; and Eastman School of Music, Rochester, New York, April, 2001.
- MOUNTAIN, Rosemary, *An Investigation of Periodicity in Music, with reference to three 20th-century compositions: Bartók's Music for Strings, Percussion and Celesta, Lutoslawski's Concerto for Orchestra, and Ligeti's Chamber Concerto*, Ann Arbor, Mich, UMI, 1993.