Perspectives on the Design of Musical Auditory Interfaces

Grégory Leplâtre and Stephen A. Brewster
Department of Computing Science
University of Glasgow
Glasgow, UK
Tel: (+44) 0141 339 8855
Fax: (+44) 0141 330 4913
Email: {gregory,stephen}@dcs.gla.ac.uk
Web: http://www.dcs.gla.ac.uk/~{gregory,stephen}

Abstract
This paper addresses the issue of music as a communication medium in auditory human-computer interfaces. So far, psychoacoustics has had a great influence on the development of auditory interfaces, directly and through music cognition. We suggest that a better understanding of the processes involved in the perception of actual musical excerpts should allow musical auditory interface designers to exploit the communicative potential of music. In this respect, we argue that the real advantage of music as a communication medium relies on the richness of its specifically musical meanings rather than on its formal structure. Finally, we propose a method for automating the design of musical auditory interfaces, in order to make this design possible to non-musician designers.
Keywords: human-computer interaction, auditory interfaces, music perception

1 Introduction

This paper addresses the use of music as a communication medium in auditory display. The function of auditory display is to help a user monitor and comprehend whatever it is that the sound output represents in a human/machine interface (Kramer, 1994). For instance, sounds can be used in a computer-based system to provide users with various information, in addition or as an alternative to graphical information. The potential of music as a medium of communication has been highlighted by various authors (Blattner and Greenberg, 1992, Smoliar, 1994, Alty, 1995), but so far this potential has not been fully exploited (Gaver, 1997).

Investigation methods in music cognition have been influenced by psychoacoustics and psycholinguistics in the last two decades (Aiello, 1994a). We suggest that a similar influence has played a significant role in the field of auditory display. Partly, because psychoacoustics provides us with perceptually pertinent low-level parameters to control the auditory medium, and because music cognition research is used as the foundation for using music in auditory interfaces. We will highlight the fact that there is a similar division between two approaches of music cognition and two approaches for using music in auditory display. Roughly, one can assert that the first approach envisages music with low-level variables, as opposed to a global perspective upon music.

Consequently to the fact that the control of low-level parameters such as pitch, is used to represent an informational structure, the attempts to use music as a communication medium mostly consist so far in mapping the informational structure with the structure of a musical sentence. Since the communication of information relies in the understanding of the structure of a musical excerpt rather than in its actual musical *meaning*, one can argue that such an auditory representation language does not take advantage of the specific power of the medium, which is its musicality. We believe that the progresses of music cognition in understanding how music is globally perceived should help the auditory interface designer to use musical meanings effectively within a musical representation language.

One of the major interests in the research into auditory display is to formulate design principles that can be used by non-experts. Still, it seems difficult to provide non-musician designers with guidelines that would allow them to create musical materials. We propose that the automation of the design process should not require the designer to have musical knowledge. In an example, we show that the interaction process can be specified by a set of rules. The musical expertise needed would be restricted to the development of the system allowing any designer to use music effectively in human/machine interaction

2 Music: a Potentially Powerful Medium of Communication

Ever since the audio channel has been used to convey information in human/machine interfaces, various authors have addressed the issue of using music for this purpose. Blattner and Greenberg (1992) write:

> "Music has a communicative aspect not limited to the absolutes of spoken language. Additionally, the "emotional" responses of music, subjective though they may be, can, if harnessed properly, be of tremendous import to the transmission of non-speech audio information".

Smoliar (1994) points out that when we need a communication medium that involves more than the exchange of words, music is one of the better known disciplines that communicates powerfully through non-verbal means. He argues that since communication is an act of intelligent behaviour, looking at music, rather than natural language, we can more clearly focus this vision of communication as a behavioural process. Alty (1995) also highlights the potential of music as a communication medium:

> "Music is all-pervasive in life and forms a large part of people's daily lives. It is very memorable and durable. Most people are reasonably familiar with the language of music in their own culture. Once learned, tunes are difficult to forget".

There is a long tradition of communicating through non-speech sound like music, horns and bells in Europe:

> "Hunting horns are an excellent example of signal type non-speech messages (…). These messages included warnings, cheering on the hounds, call for aid, fanfares for each animal, and so on" (Blattner and Greenberg, 1992).

And drums in Africa:

> "surely one of the most remarkable methods of communication is the talking drum of central Africa (…). The languages spoken in the areas of central Africa where the talking drums evolved are pitched. There are two tones, high and low, that are used variously with each syllable of a word. The talking drums also have tones, high and low, which imitate the tonal patterns of words" (Blattner and Greenberg, 1992).

Earcons are structured non-speech sounds that can be combined, transformed, inherit from other earcons properties and constitute an auditory language of representation (see Blattner *et al.* (1989) for an introduction to earcons and Brewster *et al.* (1993) for more up to date information). Blattner and Greenberg (1992) suggest that these messages should play the role of chorus in Japanese Noh drama (in Noh drama, a chorus is part of a coded language that transmit information about the context of the dramatic situation), though we can argue that this language does not take advantage of the specific meanings of music. Indeed, the effectiveness of earcons relies on the fact that people have to learn the structure of the sounds in which information is contained. One can argue that, on the contrary, music transmits information without requiring its structure to be understood. This will be discussed in the next section.

Since we have the technology to create any possible sound, it is possible to take advantage of the universal meaning of music to create rich soundscapes that enhance and intensify our computer interfaces. This challenge is still to be met. Accordingly, Gaver (1997) outlines in a recent study that auditory interfaces have so far drawn very little on the possibilities suggested by music. He proposes a possible explanation in that the control needed for the research on auditory interfaces implies a level of explicit articulation, which the complexity of music resists. He continues:

> "This situation contrasts with designers of multimedia or games environments, who happily exploit music's potential to create mood without needing to articulate exactly how they are doing so."

Yet, despite the theoretical potential of music to transmit information reviewed in the first section of this paper, it turns out that, so far this potential could not be fully exploited. In order to understand why music resists being used in auditory display, let us precisely define what is an auditory representation. Kramer (1994) envisages auditory display in terms of two broad categories: *analogic* and *symbolic* information representation.

- “An analogic representation is one in which there is an immediate and intrinsic correspondence between the sort of structure represented and the representation medium (…). By way of example, a Geiger counter produces an analogic representation.”
- “By symbolic representation we refer to those display schemes in which the representation involves an amalgamation of the information represented into discrete elements and the establishment of a relationship between information conveying elements that does not reflect intrinsic relationships between elements of what is being represented”

Analogic representation clearly relies on the use of auditory variables matching informational variables. In this highly constrained medium of communication it seems like meaning of auditory messages belongs entirely to its structure. As opposed to that, we believe that specific meanings of music (discussed in the following section) can be particularly useful to symbolise a piece of information. Smoliar (1994) writes that: “music is the purest form of rhetoric, because for all intents and purposes, it is totally unencumbered by either syntax or semantics”. This suggests that, indeed, one should consider using music for its specific musical meaning rather than for its constitutive structure.

3 The Inheritance of Psychoacoustics and Psycholinguistics

Although music is perceived as a high-level object, musical display is essentially controlled with low-level variables. First because this is a convenient means of control, and also because psychoacoustics and music cognition research focus on such parameters. Thus the perceptually meaningful parameters available for the designer are principally low-level. In this section, we show that there is a parallel between a class of music cognition research, in which experimental processes are inherited from psychoacoustics and psycholinguistics, and musical display, based on the use of low-level parameters. On the other hand, one can associate the cognitive approach for which listening is not passive, but an activity of individuals, to an approach of musical display for which musical sounds stand for high-level objects.

The audio channel is only controllable through variables. Low level parameters such as pitch or intensity are very convenient to use in this respect. Now psychoacoustics allows making the connection between physical structures of sounds (or sound organisation) that designer control, and the perceptive effects (representing an information structure) they are aiming for. In addition, psychoacoustics is an indispensable tool to get the best out of the huge audio possibilities available with new technologies.

On the other hand, according to Aiello (1994a), when addressing the cognitive mechanisms related to music, research was constrained within the experimental methods inherited from psychoacoustics. On this subject, Deliège and Ahmadi (1990) write:

> "The usual practice, in our field, as in any scientific discipline, is to isolate the variable that one wishes to study and to incorporate it in a series of brief and repetitive sound sequences (they are called musical), constructed by the psychologist for the need of the experiment, in order to be able to identify it afterward, in appropriate manner in the statistical analysis of the data (…). Unfortunately, many studies in the field of psychology of music scarcely achieve these aims because a musical objective is being sought through the use of material that is both too simple and too trivial."

In addition, the psychology of music has also been influenced by research in psycholinguistics as Aiello (1994a) suggests:

> "As in psycholinguistic research, the emphasis in the psychology of music has been on investigating the phonetic and the grammatical levels of brief strings of words without an in-depth look at the perception of entire discourses, so research in the cognitive psychology of music has examined the perception of brief melodies or sequences of musical stimuli, and has not thoroughly investigated the perceptual processes that take place when listening to longer musical excerpts."

In the field of cognitive science there is a division between the two following opposite approaches:

1. Firstly, a bottom-up approach. In accordance with a fragmented representation of the world, all the phenomena are represented with an input, a linear mechanism, and an output. This vision doesn't take into account the top-down mechanism of cognition. In this respect, individuals stand for receptive and contemplative.

2. In contrast with that, there is an approach for which cognitive activity is not contemplative, but active. Therefore, objects have a more specific cognitive significance than a natural meaning.

Beyond this theoretical division, there are practical questions of methodological processes. And what matters is what we can learn from both. About the bottom-up approach, Clarke and Krumhansl (1990) write:

> "There are certain obvious advantages in this very controlled kind of approach, and it has proved extremely powerful and productive for advancing our understanding of tonal and metric hierarchies. However, it has left untouched a range of issues concerned with listeners' understanding of more extended and elaborate structures in which a considerable degree of interaction between different parameters can be expected."

A great deal of work has investigated the interaction between different parameters from a low-level viewpoint. For instance, Thompson (1994) proposes three different way to study parameter interaction in music psychology. Schmuckler (1994) gives other evidence of the interdependence of different variables in his study on harmonic and rhythmic influences on musical expectancy. If these studies are providing us with knowledge of musical cognition, are they aiming at a global understanding of this activity? Aiello (1994a) argues that research in the perception of music cannot remain research in the perception of isolated musical variables (and even interactions between a limited number of these variables) couched within very brief stimuli, because music is more than the sum of its parts and is experienced over a temporal continuum.

In auditory display, there is a similar division. In reference to perception of music as perception of isolated musical variables, Brewster (1995) drew general principles for the design of earcons. In contrast with that, some designers choose to use auditory messages as high level objects. As yet, this paradigm has mainly been applied to real-life sounds (see (Gaver, 1986) for an introduction to *auditory icons*, (Mynatt *et al.*, 1998), and (Cohen, 1994a, 1994b)) but also to music (Bargar, 1994).

Moreover, auditory languages of representation need to be evaluated. Low-level variables will provide designers with convenient evaluation. Still, the experimental processes do not focus on these parameters, but on the performance of a certain task. The parameters can then be adjusted or changed to improve the performance of the evaluated task. When sounds are used as high-level objects, the evaluation of the related representation language can only tell the designer that his language is effective or is not. Then it is up to the designer's intuition and experience to make relevant modification to the sounds.

As it turns out that research in music cognition is the grounding of musical interface design, the development of research for understanding global audition of music should significantly improve the control an the efficiency of musical representation language in the field of auditory display.

4 Language, Meaning, and Musicality

Despite the fact that music is not universally envisaged as a language, music and language feature obvious similarities (Clarke, 1989b, McAdams, 1987, Sloboda, 1985). In musical display, the transmission of information is based on the use of a musical representation language. Since the language of music is greatly understood by most people, a representation system based on it would require a limited period of time to be learned. Here it matters to make clear the distinction between two distinct linguistic levels: in the same way as we talk about compositional language, we can talk about auditory representation language. Both of these notions differ from the language of music as a mind process. It is probable that a good understanding of the musical linguistic process would be helpful for designers to create pertinent representation languages. If the linguistic understanding of music and language turns out to be similar (Lerdahl and Jackendoff, 1983), it seems evident that music cannot transmit the same class of information as language does. Indeed, music mainly addresses to feelings as opposed to the language of words. Aiello (1994b) asserts:

> “Music and language are both modes of communication, yet they have different goals. Generally speaking, while the primary aim of language is to communicate thought, one of the main goals of music is to heighten emotions and express them aesthetically. Music is born out of the need to express ourselves and to communicate aesthetically through the abstractness of the characteristics of sound.”

Yet, using the emotional potential of music is not acceptable for being too intrusive in a working environment. How can we take advantage of the specificity of music as a neutral (not too intrusive) communication medium? Furthermore, is it reasonable to contemplate conveying information through music without replacing the musical substance of the medium by a soulless piece of information? What would be the point of using music, then? To address this issue, let us refer first to Meyer’s thesis related to meaning in music. To the question: Where does the meaning lie in music, Meyer (1956, 1994) answers with the following distinction: for “absolutists”, “the meaning of music lies exclusively within the context of the work itself, in the perception of the relationship set forth within the musical work of art”. On the contrary, “referentialists” insist that “in addition to these abstract, intellectual meanings, music also communicates meanings which in some way refer to the extramusical world of concepts, actions, emotional states, and character.” We believe that, in the field of musical display, the meaning of music belongs even less to music than from the refererentialist point of view. In this particular situation of listening, one could call this act of listening *contextual*. In many cases, indeed, audio information is processed by users in the background of their main task. But what mostly justifies the definition of this notion of contextuality is the very specific functional status of music in an interaction situation. In this respect, there are three levels of user’s music understanding that should be taken into consideration:

- The *intellectual* level. At this level, a meaning lying strictly in the musical work of art arises.
- The *referential* level. An additional meaning is related to the extramusical world of concepts, actions, emotional states, and character.
- The *contextual* level. Music takes its actual informational meaning at this level. This is specific to a situation of interaction in auditory display.

What is noticeable here is that the last level is the one responsible for musical objects to be actually transformed into a non-musical piece of information. We assert that the high degree of functionality associated with music in an interactive situation is playing against the perceived musicality of the display.

In an informal discussion, George Bloch, a French composer, acknowledged that, as a professional musician, the occasion when he could enjoy a piece of music were precious, since his analytical listening would almost inevitably take over the "natural way" of listening to music. This anecdote shows that as in auditory display, when analytical audition takes over, it loses a part of its musicality. Hence, if a reasonable amount of musicality needs to be preserved, non-informational musical elements should be preserved within the informational audio structure.

It is obvious here that musicality is not only part of music but an activity of individuals listening to it. Alty (1995) writes: "The information contained in a large scale musical work (say a symphony) is very large (a typical audio CD contains many hundreds of megabytes). The information is highly organised into complex structures and sub-structures. The potential therefore exists for using music to successfully transmit complex information to a user." Again, assuming that the complex structures embodied by music could provide a mapping for complex informational structures means that the user is able to understand how a piece of music is constructed. Now if he/she does, it is at the expense of the perceived musicality of the tune.

The main reason why musicality is an issue in auditory display is that it is related to annoyance. We believe that a musical excerpt does not sound annoying in a working environment if it is reasonably pleasant and neutral. But this does not mean that non-speech audio sound should be musical to be pleasant. It has to meet users' expectancies. Thus, if a system introduces musical features, they ought to sound pleasant. On the other hand, if sounds are structured without involving musical functions, experiments have shown that the sounds are not annoying at all, according to Brewster *et al.* (1996). Again, it is related to users expectancies. As a comparison, listening to a loud bell 12 times when we want to know that it is time for lunch would not sound annoying at all since it has an expected function. Eventually we do not mind counting the gongs. As long as the tune has a function, its musicality does not matter. It just stands for one of the numerous informational audio features of or environment.

On the other hand, if a communication sound embodies musical features, (e.g. features that recall a musical experience to individuals) the more people will become familiar with it and its signification, the more they will listen to it as a musical object. If the tune sounds unpleasant, it will become rapidly annoying.

From a more practical point of view, a little effort in the actual creation of musical materials could improve their pleasantness a lot. Technology allows us to create any possible sound, but the downside of it is the ubiquity of low quality imitations of real instruments in computer-based systems. These mortifying periodic sounds, as Edgar Varèse called them, play at the expense of the pleasantness of the interface. But using a good synthesiser is not sufficient to provide musicality to an interface. Another drawback of technology is related to interpretation of music. Richness of interpretation can still be transmitted when the material used is pre-recorded, but it is more problematic when it must be synthesised in real-time by a machine. Interpretation is indeed fundamental in the communication process according to Lewers (1980): "If inflection and nuance *enhance* the effect of spoken language, in music they *create* the meanings of the notes". In conclusion, particular attention should be paid so that Cook's (1994) claim, "Perceptual psychologists assume that music is made of notes" does not apply to auditory interface designers.

Thus, it is now clear that the particular meaning of music, that goes far beyond its structure, has to be wisely considered by auditory interface designers for them to get the best out of the communicative potential of music.

5 Perspective: Towards Intelligent Design

In this section, we present a specific example of auditory representation and a perspective for automating this representation. The background of this research is improving interaction in telephone-based interfaces, such as mobile phones. In non-visual interaction, Yankelovich *et al.* (1995) suggest that navigating in hierarchical menu is a major problem. Brewster *et al.* (1996) showed that earcons could be used to help users navigating in such structures.

Using earcons to provide navigational cues has been proved to be effective in previous research. Earcons are constructed from *motives*. Motives can be composed of a single note or a group of notes, but are always an elementary object of the representation language. By modifying the psychophysical parameters of the sounds from which motives are constructed, it is possible to create hierarchical earcons. With the idea of automating the auditory display, we attempt here to give a model for this particular design task.

A hierarchical structure can be represented by a tree; let us use the following numerical denomination of the tree nodes suggested by the following graph:

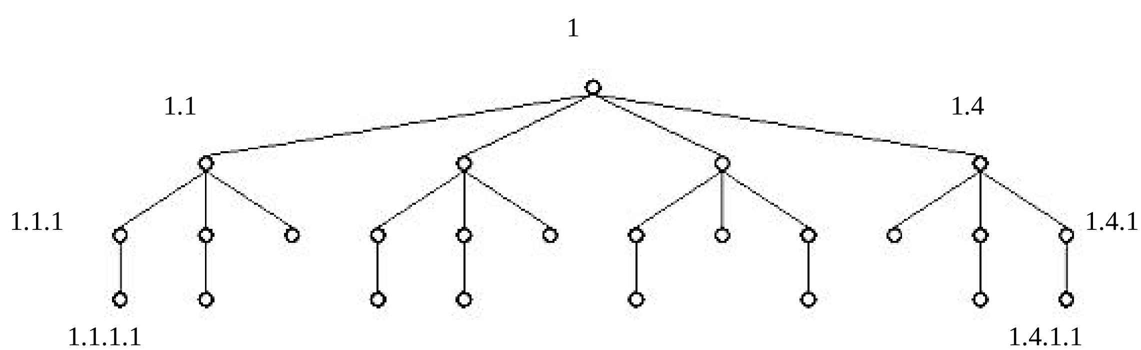

Figure 1: Hierarchical Structure Investigated

This particular tree (Figure 1) has been used in several experiments to evaluate the effectiveness of different audio representation based on earcons (see Brewster *et al.* 1996).

Here we can choose between two main categories of meanings in each sound:

- An absolute meaning. The motive has to be recalled without any reference to the other motives. It can be a chord, or a melody, or a more structured tune.
- A referential meaning. A motive is associate with another motive with a logical link. For example, a motive that is the repetition of another motive.

The balance between these meanings in the hierarchical structure is fundamental. Indeed, too many motives with absolute meaning would make the set too hard to remember. On the other hand, to many logical links between a restricted amount of motives would require the user to focus on structural features of the sounds at the expense of their perceived musicality, according to the discussion of the previous section.

In order to formulate the design process, we need to assume the availability of different functions (There is no strong theoretical grounding for these functions, but they are convenient to give a formal description of the design process):

- P, a perceptive measure of a motive "quality". Of course, the values of P can hardly be numerical. The idea is just that motives should be classified in a perceptually pertinent way.
- d, a measure of the perceptual distance between two motives.

The design task can be specified with the preceding functions as follows (the root is at level 0):

Level 1:
The sounds of this level are chosen to be as distinct as possible since they will be characteristic for each of the four associated sub-trees. This can be easily formulated as follows:

Maximise $d(P(1.1),P(1.2))$
Maximise $d(P(1.1),P(1.3))$
Maximise $d(P(1.1),P(1.4))$
Maximise $d(P(1.2),P(1.3))$
Maximise $d(P(1.2),P(1.4))$
Maximise $d(P(1.3),P(1.4))$

Level 2:
At this level, the sounds must be understood as deriving from the sound of their parent, and they are also chosen to be related to each other within this level. For example, 1.1.1, 1.1.2, and 1.1.3 inherit from the perceptual properties of 1.1, and are all linked with a logical relation that we note R:

$P(\mathrm{s}.1)=P(\mathrm{s})+\mathrm{p}_1$

$P(s.2)=P(s)+p_2$
$P(s.3)=P(s)+p_3$, for s=1.1, 1.2, 1.3, 1.4

And:

$p_1 \; R \; p_2 \; R \; p_3$

In addition, to make the set easier to remember, the logical relation used at level 2, within each of the four sub-trees, can be the same for all the sub-trees For example, if within one sub-tree p1 is represented by a chord, p_2 by a sequence of 2 chords, and p_3 by a sequence of 3 chords, within another sub-tree, the same principle can be used. For both sub-trees, the sequences of chords will match the musical context determined by the motive of root at level 1. This can be formulated as follows:

$P(s.1)=P(s)+p_1$
$P(s.2)=P(s)+p_2$
$P(s.3)=P(s)+p_3$

$P(t.1)=P(t)+q_1$
$P(t.2)=P(t)+q_2$
$P(t.3)=P(t)+q_3$, for s,t=1.1, 1.2, 1.3 ,1.4

And:

$p_1 \; R \; p_2 \; R \; p_3$
$q_1 \; R \; q_2 \; R \; q_3$, where R is a single relation.

Level 3:
The rules for this level are chosen to be similar to the level above. So we can write:

$P(s.1)=P(s)+p$, where s is any node of level 2.

We can also impose the inheritance process from level 2 to level 3 to be the same for all level 2 nodes with this constraint. For example, All level 3 earcons :can be constituted of the sequence of the corresponding level 2 earcon with a chord.

$(P(s.1)=P(s)+p)$ R $(P(t.1)=P(t)+q)$, for all level 2 nodes s,t.

Of course, it only makes sense to construct a rule-based model in the perspective of solving it. Typically, this system of rules can be implemented in a language like Prolog, but its solution is conditional to the existence of the following features:

- A base of solution, containing the classified musical materials.

- Additional rules, according to which the different motives can be transformed and structured together into coherent musical sentences.
- Rules related to the tasks earcons have to support.

We have implemented so far a set of hierarchical sounds involving musical features by solving this system “empirically”. This set of sounds has been evaluated and the results of the experiment showed that participants could recall the sounds successfully after a short training. We can now consider automating the design process confidently. On the other hand, the tasks actually performed while navigating in a hierarchical structure still have to be examined since these tasks shall have an influence on the design of the hierarchical earcons.

The important point here, as regards music cognition is that the creation of the latter rules does not only require a musical knowledge, but also a good knowledge of the perception of formal structures involved in music. Yet a decade ago, Clarke (1989a) was pointing out the gap between formal structures and their psychological knowledge. Research performed in this area in the last ten years should be extremely relevant for the future development of a system that would allow non-musician designers to produce effective and efficient musical interfaces

6 Conclusion

This paper proposed a point of view on the use of music in auditory human-computer interaction. The aim of this study was to highlight the specific status of music within the audio channel in order to make a good use of it as a communication medium. We showed that the influence of psycholinguistics and mostly psychoacoustics on music cognition research plays a fundamental role in the design of musical audio interfaces. Since musical sentences are perceived as high-level objects, it is probable that the development of music cognition of global pieces of music should help musical interface designers to exploit music for its global meaning rather than for its structure.

Therefore, one must focus on the actual meaning of musical sentences on which one can rely to convey information. What makes the specificity of a musical sentence is its musicality. Yet, one can argue that this musicality is not inherent to the sentence, but depends on the informational content the designer allocated to it. As a consequence, it seems that a reasonable approach to a musical design consist in balancing purely musical elements and musically informational elements within the representation language.

Furthermore, we believe that automating the design process is a most interesting issue since it could allow non-musician designers to build auditory interfaces that take advantage of the potential of music. We have shown that the automation of this process could be performed easily by using a rule-based specification language. Again, the implementation of such an automatic system relies on a deep understanding of the mechanisms involved in the cognition of complex pieces of music.

Acknowledgements

This research was founded by EPSRC; grant number GR/L66373.

References

Aiello, R. (1994a). Can listening to music be experimentally studied? In Rita Aiello with John A. Sloboda, editor, *Musical Perceptions*. Oxford University Press.

Aiello, R (1994b). Music and language: Parallels and contrasts. In Rita Aiello with John A. Sloboda, editor, *Musical Perceptions*. Oxford University Press.

Alty, J. L. (1995). Can we use music in computer-human communication. In *Proceedings of HCI'95*, pages 409-423.

Bargar, R. (1994). Pattern and reference in auditory display. In Gregory Kramer, editor, *Auditory Display*, Addison-Wesley.

Blattner, M. M., and Greenberg, R. M. (1992). Communicating and learning through non-speech audio. In Edwards, A. and Holland, S., editors, *Multimedia Interface Design in Education*, pages 133-143. Springer-Verlag.

Blattner, M. M., Sumikawa, D. A., and Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human-Computer Interaction*.

Brewter, S. A., Wright, P. C. and Edwards, A. D. N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of INTERCHI'93* (Amsterdam, Holland) ACM Press, Addison-Wesley, pages 222-227.

Brewster, S. A., Raty, V.-P., and Kortekangas, A. (1996). Using earcons to provide navigational cues in a complex menu hierarchy. Technical report, Department of Computing Science, University of Glasgow.

Brewster, S. A., Wright, P. C., and Edwards, A. D. N. (1995). Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of BCS HCI'95,* pages 155-159, Huddersfield, UK.

Clarke, E. F. (1989a). Mind the gap: formal structures and psychological processes in music. *Contemporary Music Review*, 3(1):1-13.

Clarke, E. F. (1989b). Issues in language and music. *Contemporary Music Review*, 4: 9-22.

Clarke, E. F., and Krumhansl, C. (1990). Perceiving musical time. *Music Perception*, 7:213-251.

Cohen, J. (1994a). Monitoring background activities. In Kramer, G., editor, *Auditory Display*. Addison-Wesley, pages 499-531.

Cohen, J. (1994b). Out to lunch: Further adventures monitoring background activitiy. In Kramer, G., and Smith, S., editor, *Proceedings of the second International Conference on Auditory Display*, Santa Fe, pages 15-20.

Cook, N. (1994). Perception: a perspective from music theory. In Rita Aiello with John A. Sloboda, editor, *Musical Perceptions*. Oxford University Press.

Deliège, I., and El Ahmadi, A. (1990). Mechanisms of cue extraction in musical groupings: A study of perception of *Squenza IV* for solo viola by Luciano Berio. *Psychology of Music*, 18(1):18-44.

Gaver, W. W. (1986). Auditory icons: Using sound in interfaces. *Human-Computer Interaction*.

Gaver, W. W. (1997). Auditory interfaces. In M. G. Helander, T. K. L. and Prabhu, P., editors, *Handbook of Human-computer Interaction*. Elsevier Science, Amsterdam, second edition.

Kramer G. (1994). An introduction to auditory display. *Auditory Display*. Gregory Kramer Editor, Addision-Wesley.

Lerdahl, F., and Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA, MIT Press.

Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press.

Meyer, L. B. (1994). Emotion and Meaning in music In Rita Aiello with John A. Sloboda, editor, *Musical Perceptions*. Oxford University Press.

McAdams, S. (1987). Music: A science of the mind? *Contemporary Music Review*, 2: 1-61.

Mynatt, E. D., Black, M., Want, R., Baer, M., and Ellis, J. B. (1998). Designing audio aura. In *Proceedings of CHI'98*, ACM, Los Angeles, pages 866-573. Addison-Wesley.

Nettl, B. (1983). *The Study of Ethnomusicology*. University of Chicago Press, Urbana.

Schmuckler, A. M. (1994). Harmonic and rhythmic influences on musical expectancies. *Perception and Psychoacoustics*. 56(3):313-325.

Sloboda, J. A. (1985). The musical mind. Oxford University Press.

Smoliar, W. (1994). The role of music in multimedia. *IEEE Multimedia*, pages 9-11.

Thompson, W. F. (1994). Sensitivity to combinations of musical parameters: Pitch with duration, and pitch pattern with durational pattern. *Perception and Psychoacoustics*, 56:363-374.

Yankelovich, N., Levow, G., and Marx, M. (1995). Designing SpeechActs: Issue in speech user interfaces. In *Proceedings of CHI'95*. Pages 369-376. ACM Press. Addison-Wesley.