

# Auditory Interfaces

# A Design Platform

Dan Gärdenfors  
gardenfors@hotmail.com  
2001

## Contents

### 1 Introduction

### 2 Background

#### 2.1. Why Auditory Interfaces?

#### 2.2 Hearing and Vision

#### 2.3 The Potentials of Auditory Interfaces

### 3 Mapping Sound to Information

#### 3.1 Speech-based Interfaces

#### 3.2 Iconic, Symbolic and Metaphoric Interfaces

### 4 The multidimensionality of sound

#### 4.1 The Fundamental Dimensions of Sound

#### 4.2 Manipulation of Sound Parameters

#### 4.3 Musical Interfaces

### 5 Conclusions

## Reference List

# 1 Introduction

How can sound feedback improve human-machine interaction? This essay aims to serve as a theoretical introduction to auditory interface design. Even though hearing and vision are our two primary senses, most interfaces are today mainly visual. As vision and hearing are fundamentally different, there are several advantages to using auditory feedback in interfaces. These are introduced in the first part of the study. Next, three different strategies for how sound can be used in human-machine interfaces strategies are compared: speech, iconic sounds and symbolic sounds. However, since speech as a medium for communication is very well understood, emphasis is on auditory interfaces that mainly rely on non-speech sounds. Furthermore, there are several disadvantages to only relying on speech in an interface. This study aims to evaluate the three key approaches with emphasis on efficiency and aesthetics. Finally the palette that is available to designers of auditory interfaces is introduced by describing sound as a multidimensional medium. The discourse deals with sound on a fundamental scale, followed by a discussion of the potentials of larger musical interfaces.

The main sources cited in this essay are articles on human-computer interaction and ergonomics. Particularly a chapter by William W. Gaver in the Handbook of Human-Computer Interaction (1997) has been used. Gaver's text summarises experimental findings on auditory interfaces, results stemming from original reports that are hard to retrieve. Another paper by Gaver, The Sonic Finder (1989), provides central information to the use of everyday sound in auditory interfaces. Stephen A. Brewster presents an inspiring study of "earcons" in his paper Using Non-speech Sound to Provide Navigation Cues (1998).

The field of auditory interfaces is rather new, with much yet to be explored. Because of this, many arguments presented here are mainly hypothetical, and since no sounds are provided, they assume that the reader has a good auditory imagination. This essay is based on the assumption that most people can differentiate between sounds if they are different enough. I believe people can interpret sounds fairly well, even if only few might have the vocabulary to express their understanding, as experiments by Brewster (1998) indicates. However, while empirical findings are crucial parts of psychological generalisation and usability issues, there are still very few experimental conclusions in the field of auditory interfaces. Nevertheless, there seems to be some general traits that characterise sounds that are useful auditory interface material. The ideas put forth in this essay are mainly to function as general hints. Obviously, if one is to develop a good interface, much specific research and testing is

necessary. Even within the guidelines of this essay, considerable amounts of creativity are needed to design a useful and aesthetic auditory interface.

## 2 Background

### 2.1 Why Auditory Interfaces?

Interfaces for computers, mobile phones and other machines today mainly present information visually. While it often is convenient to display information on monitors, visual interfaces also have crucial limitations. One issue is that displays are very size-dependent, since size corresponds closely to the amount of information that can be conveyed. Size always seems to count when developing portable electronic devices, since most people prefer to carry the smallest and lightest machines possible. One efficient way to decrease the size of a portable device is to decrease the size of the visual display or - more radically - completely remove it. This is possible if information can be conveyed aurally instead of visually. Auditory interaction is necessary, for example, when people want to communicate with computers through a telephone. According to Brewster (1998), telephone-based interfaces are becoming increasingly important for human-machine communication. Examples of widely used telephone-based interfaces are booking tickets or performing bank transactions over the phone.

Another problem of visual displays is that the user must focus on it to obtain information. Auditory feedback, however, enables the user to look away from the device he or she is using. Consequently, the user may be able to perform more than one task at a time, such as driving a car while using a telephone or grabbing a cup of coffee while waiting for a computer to finish downloading a file. Auditory feedback can often be a necessary complement, but also a useful alternative to visual feedback. When designing a mobile electronic device, it is difficult to predict all possible scenarios when it might be used. Obviously, visual feedback is preferred in many situations such as in noisy environments or when the user has to concentrate on a

listening task. However, as there might be numerous occasions when a user cannot look at a display, versatile devices such as mobile phones or hand-held computers benefit from having flexible interfaces. Still, it seems that auditory feedback has been overlooked in many common human-machine interfaces.

## 2.2 Hearing and Vision

Vision and hearing are our two primary senses for obtaining information about the outside world. Hearing has often been considered secondary to vision, as it seems that in many situations we use our ears merely to tell us where to turn our eyes (Gaver 1997). However, it is important to emphasise that sound is a unique medium that can provide information which vision cannot.

Our eyes perceive light, which is reflected from objects around us. Vision hence tells us about the surface, size and shape of objects. Our ears, on the other hand, perceive patterns of moving air that vibrating objects generate. Sound can carry information about the consistency and hollowness of objects. Hearing can therefore provide understanding about the interior of objects, which is a domain where vision is limited. Another feature of sound is that it can communicate information quickly (Brewster 1998). Sound is of a fundamentally different temporal nature to that of visual objects; what we hear to be more transitory than what we see. In the words of Gaver (1989), "sound exists in time and over space, vision exists in space and over time." Spatially, sound has the advantage of not being bound to a certain location. To see something, say a computer screen, we need to face it. However, the sound from a speaker can be heard in darkness, from far away and facing any direction. A drawback of this is that one cannot turn away from sounds. Neither can one close one's ears from an unpleasant sound.

In our everyday lives, sound and vision interact smoothly. Hearing and vision complement one another in the natural world around us and could also do so in films, multimedia and other environments created by human beings. People prefer to communicate face to face, being able emphasise words with facial expressions and body language. Naturally, almost every form of communication – e-mail, letters or even talking on the phone – has characteristic limitations. For example, written words cannot convey intonation as well as a spoken voice. Human-machine interaction ought to benefit from using sound because it is central to human communication. If the possibility of conveying information sonically were used to its full potential, it would be a powerful complement to visual interfaces. A strong argument against the use of sound in interfaces is that it easily can

become annoying, since it is more intrusive than visual impressions. However, by skilfully designing auditory interfaces, this can be avoided.

### 2.3 The Potentials of Auditory Interfaces

As many interactive systems today provide no sound feedback at all, present auditory interfaces are usually limited to a couple of warning sounds that are emitted in the case of extreme events. Sound feedback provided in mobile phones or other small electronic devices are often mainly simple beeps. Other common auditory interfaces are those in the operative systems of personal computers. However, the sound feedback in MacOS or Windows, is provided only on a few occasions. The most advanced auditory feedback seems to exist in computer games and multimedia products. Gaver (1997) claims that memory limitations in the technical product is one reason why sound feedback has not been used on a larger scale. Until quite recently it has been too expensive computationally to use sound of good quality in computers. Today, only lightweight electronic devices, such as mobile phones or hand held computers have limited memory capacities, although this is rapidly changing with the development of memory cards and effective compression algorithms for sound. Seeing that the potential to use sound in electronics is growing fast, Donald Norman (1990) claims that the use of sound based interfaces is only in its infancy.

Auditory feedback in existing interfaces is commonly limited to various kinds of signals. Typical signals are sounds that indicate some sort of warning or alert, such as alarm sounds and low battery level warnings. Other signals provide feedback that some event has been successful, such as when buttons are pressed or machines are switched on. Yet, there are several types of events that have not been commonly associated with sound. For example, Brewster (1998) has investigated the use of sound as a provider of navigational cues in menu hierarchies. Such menu hierarchies are common structures in computers, mobile phones or telephone-based interfaces. Gaver (1997) claims that another little explored area is to use sound for communicating ongoing processes. There are innumerable examples of real life events that emit sounds while active: water flowing, food frying, a fan humming, a disk drive whirring and so on. The auditory interface of operative systems would grow extensively if ongoing processes were to generate sound. If using continuous sounds as opposed to the more common brief signals, auditory interfaces do not need to be more transitory than visual interfaces. However, such sounds probably benefit from being quite discreet. While, most existing sound feedback today occurs in the foreground of the

interface, subtle background sounds can be a useful complement in advanced auditory interfaces. Films and computer games generally make use of music and sound effects smoothly, so that they do not interfere with the visual information conveyed. In films particularly, music is often effectively used to enhance the visual impressions by manipulating the viewers' mood. In a similar way, an interface that uses sound cleverly can enhance the user's immersion and improve interaction. Gaver (1997) found that during an experimental process control task, the participants' engagement increased when provided relevant sound feedback. By developing more efficient auditory interfaces, interaction with machines can become easier, and hopefully more pleasant.

As there are many ways in which sound can be employed in interfaces, it is important to define the purposes of every sound at an early stage in the design process. A sound that conveys crucial information should have different attributes to one that serves as a complement to visual information. It is important to distinguish between these two very different approaches. Therefore, I choose to call them the practical and the naturalistic approach to sound feedback. The "practical" approach to auditory interfaces deals with sound as the main feedback. This can be the case when designing interfaces for visually impaired people, who must rely on sound feedback to provide sufficient assistance in performing a task. Sound feedback is also crucial in telephone-based interfaces, for example when executing bank transactions over the phone. Furthermore, sound is often the only means of communication when using a portable handsfree device with a mobile phone. Auditory interfaces based on a practical approach should be comprehensive and simple. The drawback of using very apprehensive sounds is that they might be noisy and tiresome over time.

The "naturalistic" view regards sound mainly as a complement to a visual interface. A naturalistic interface combines sound and vision in a way as similar as possible to corresponding phenomena in the natural world. Such auditory interfaces are supposed to enhance interaction between the user and a machine, especially in situations where the visual interface is ineffective on its own. Sounds that complement a visual interface can generally be subtle background events that do not disturb. In a way, such sounds correspond to the background music of films, since they convey information to the audience without interfering with the main events. Sound feedback based on the naturalistic strategy is thus very subtle and might only be recognised subconsciously. Passive learning and conditioning are interesting effects by subtle stimulation that need to be further investigated if using such feedback.

# 3 Mapping Sound to Information

## 3.1 Speech-based Interfaces

Speech is the most obvious form of information-carrying sound; spoken language is the predominant form of human communication. Moreover, speech is a very specific way to convey auditory information. At a first glance, it seems convenient to base an auditory interface on recorded or synthesised speech. However, there are several reasons not to convey all auditory information verbally.

Generally, speech is slow. Listening to a voice in a mobile phone interface makes a task such as checking battery levels much slower than it could be. Listening to a recorded voice repeatedly can also be tiresome for aesthetic reasons. Speech is not always suitable to accompany ongoing processes. When copying a file on a computer no one would be interested in hearing “reading, reading, reading, reading - writing, writing, writing, writing.” Instead, maybe a subtle percussion rhythm could be heard. Since speech is always rather obtrusive and attention demanding, it should be avoided as routine task feedback if possible.

There are situations when verbal feedback is downright disadvantageous. When listening to a voice through a telephone based interface or reading a text, additional verbal information might interfere with the task. Brewster (1998) argues that navigation problems in telephone-based interfaces can occur when speech is used both to provide information and perform navigation. Still, speech is a very useful way to convey information, especially in “practical” auditory interfaces. Speech can readily be used in combination with non-verbal information. A possible application would be to use brief abstract sounds as immediate feedback; then, if the interface user gets stuck, a voice prompt could be provided on request or after a delay.

## 3.2 Iconic, Symbolic and Metaphoric Interfaces

Gaver (1997) examines two different strategies for using sound to convey information in non-verbal auditory interfaces. One possibility is to base the interface on sounds, the origins of which are analogous to what they are to represent. This correspondence characterises auditory icons, which are usually based on everyday sounds. The contrary strategy is to use sounds that are arbitrarily linked to what they represent. Earcons is a commonly used name for these symbolic sounds. An “earcon” is hence often a musical sound that can be created from any sound source. A compromise between the iconic and symbolic strategies produces metaphoric sounds, meaning that they



share some abstract feature with what they are to represent. Brewster (1998) uses the word "representational" to describe such sounds. An example of a sound metaphor is to use high pitch to represent a spatially high position. However, using low sound pressure to represent a far distance would be more of an iconic relationship, since this is what actually happens in the real world. The boundaries between iconic, metaphoric and symbolic sounds are easily blurred.

There are several advantages to using auditory icons in sound interfaces. Iconic sounds are readily learned and remembered, although they might not be altogether intuitive, as Gaver (1997) points out. When combining visual and auditory interfaces, the auditory icons can rely on the same analogy as their graphical counterparts. This is an ideal strategy, which probably produces the most comprehensive hybrid interfaces possible. Auditory icons can also be grouped to create "feedback families." An example of grouping can be found in the Sonic Finder for Macintosh computers, which was created by Gaver (1989). This operative system uses "wooden" sounds to indicate operations on text files. Hence, selecting a text file makes a tapping sound, destroying it would sound splintering and so on. Another possibility of ordering auditory icons is parameterising, where one sound quality or "parameter" corresponds to a feature of the manipulated objects. In the Sonic Finder, the pitch of the text file sounds is "parameterised" to indicate size, so that selecting a large file makes a tapping sound with lower pitch than selecting a smaller one.

A problem with auditory icons is that it can be difficult to find suitable iconic sounds for all events in an interface, since they might not correspond to a sound-producing event in the real world. (One can easily imagine similar difficulties arising if one was to create a strictly onomatopoeic language.) If compromising and only partly relying on iconic mappings, the interfaces could become quite inconsistent. Norman (1990) stresses the importance of developing a conceptual model that is understandable to the user. Hence, a well-designed interface should use as few interpretation strategies as possible. Another problem of auditory icons is that sounds from the most realistic iconic mappings might not always be suitable. A user might easily confuse an auditory interface consisting of everyday sounds with background events or noise. It can also be difficult to find sufficiently varied sounds that do not interfere with one another.

Symbolic sounds, on the other hand, are arbitrarily mapped to what they represent and do not have to be limited by any similarities. Instead, every "earcon" can be designed with emphasis on its aesthetic properties, which is difficult when developing auditory icons. Since symbolic sounds can be designed freely, it is possible to design a "musical" auditory interface that is more pleasant and less tiresome

than one based on iconic sounds. Gaver (1997) points out that music provides sophisticated system for manipulating groups of sounds. When designing a symbolic or “musical” interface, complex information can be conveyed by sounds that are parameterised in many dimensions. A closer look at sound as a medium reveals endless possibilities for “musical” communication.

## 4 The multidimensionality of sound

### 4.1 The Fundamental Dimensions of Sound

Sound is a multidimensional medium, which allows great flexibility when designing abstract auditory interfaces. An awareness of the perceptive dimensions of sound is beneficial for developing effective auditory icons. Roughly, the fundamental dimensions of sound are pitch, timbre, loudness, duration and direction (Gaver 1997). However, it must be emphasised that these dimensions are to some extent co-dependent and do not simply correspond to physical qualities of sound.

Pitch, is a quality of sound that mainly corresponds to frequency. Yet, this concept is not always applicable, since only sounds that show regular periodicity for a noticeable amount of time will be heard as having a pitch (Wishart 1996). If mapping information to pitch, discrete pitches are of little use unless the users of the interface are musically trained and have perfect pitch. However, most people ought to notice substantial differences in pitch when comparing two sounds. It ought to be more convenient to map information to pitch in terms of intervals – changes in pitch that occur within a small duration of time. The amount of people that can recognise and sing songs indicates that most humans are sensitive to contrasts between intervals.

Timbre is the most versatile dimension of sound. Timbre is a complex function of overtone structure, harmonic content, envelope, transient attack and more. Most of these acoustic elements can be shown in plots of the spectrum of a sound, which can be generated by a Fourier analysis (Wishart 1996). The possibility to create different timbres is practically unlimited. Due to this complexity, discrete timbres are easily recognisable without any need for references. Common timbres, like that of a trumpet, are readily recognised and can be remembered over time.

Loudness is a less useful parameter for communicating information. The interface user might want to be able to control the overall sound

volume of the device, which would obscure information conveyed by loudness. Moreover, distinguishing between discrete values in loudness can be very difficult. Nevertheless, as with pitch, variations in loudness over a short duration of time might be noticeable. For example, "earcons" that fade in could indicate a function being turned on and vice versa.

Differences in duration add a fourth dimension. Indeed, this dimension is intrinsic to sound. Timing is also an important feature of intervals and fades, as mentioned above. As with pitch and loudness, differences in duration can be difficult to distinguish unless they are very obvious.

Direction can be a very useful parameter when designing auditory interfaces. However, many common electronic devices, like cellular phones, are monaural. Still, if using a stereo headset for the auditory interface output, direction becomes an option. This opens up excellent possibilities for distinguishing between different sounds, as has been shown by Brewster (1998). By manipulating direction by using stereo and surround sound the dimensions of space can efficiently be mapped to information.

#### 4.2 Manipulation of Sound Parameters

Experiments carried out by Brewster (1998) illustrate the potentials of mapping information to different parameters of sound. His experiments showed that earcons is a powerful method of communicating navigation cues in telephone-based interfaces. The earcons were designed so that changes in timbre, register, spatial location and rhythm mapped to different levels and positions in the menu hierarchy. The results showed that in about 80 percent of the tests, the test persons successfully used earcons as a navigation aid. Another example of sound parameter manipulation is data auralisation. Data auralisation is achieved by mapping data variables to different parameters of sound. This can be a useful alternative to mapping data graphically, since several dimensions can be mapped simultaneously when relying on sound. Data auralisation has in several cases proven to be interpreted more accurately than graphical mappings (Gaver 1997).

By simple variations of basic sound parameters, modulations and effects such as tremolo, vibrato, echo and reverberation can be obtained. Some of these effects have useful representational properties. For example, subtle aspects of a sound can contribute to a listener's experience of the distance to the sound source. Obviously, the loudness of a sound corresponds to our perception of distance. A

subtler, but very important factor is the ratio of reverberant to direct sound (Gaver 1997). If a listener is close to a sound source, the levels of direct sound are large compared to those of echoes and reverberation. With increased distance, direct sound grows less dominant and reverberation makes up a greater part of the total sound. Also the spectrum of a sound contributes to perceived distance, since the high frequencies of a sound decrease as it travels through air. Metaphoric and iconic aspects of sound can be useful tools for many different sound applications. Frequency has been used as a representation for size, as in the parameterising of auditory icons in the Sonic Finder (Gaver 1989). When selecting a text file with the mouse a thumping sound is heard. A larger file generates a sound at a lower pitch than that of smaller files. This is an type of iconic mapping of the physical world, since natural objects usually follow the same pattern.

As mentioned earlier, timbre is a very recognisable sound parameter. Brewster (1998) used timbre as the first level organising parameter in his navigation experiment. Timbres from various common instruments can potentially also be used for their different metaphoric features. Particularly in films or multimedia productions, the timbres of several musical instruments seem to be associated with particular moods and attributes. Thus violins often represent sentimental or romantic moods; saxophones are jazzy and sexy; harps give religious associations and so on. The reasons for such associations are not always obvious, since they have often been established over a long time span. Gaver (1997) calls sounds with commonly established associations "genre sounds."

Another metaphorical aspect of timbre is how some sounds can be interpreted as more urgent than others. Factors that contribute to urgency are inharmonic timbres and abrupt onsets (Gaver 1997). Still, the expressive possibilities of sound stretch much further than to what can be achieved by means of individual sounds, as has been described above. If allowing sounds to expand over time, signals can be varied, mixed and linked in endless ways, creating advanced "languages" within an interface. When combining intervals over time one creates melodies. Longer durations of time allow sounds to be combined into rhythmic patterns. All psychoacoustic parameters mentioned above can interact, creating complex sound patterns that are best analysed in musical terms.

### 4.3 Musical Interfaces

Gaver (1997) states that focusing on attributes of sound – its pitch, loudness, duration or timbre – characterises musical listening, as compared to determining the source of a sound, which typifies everyday listening. Music is a complex phenomenon and our understanding of it is limited; there is for example not even a general and widely accepted definition of music. Nevertheless, several characteristics of Western music can be described in contemporary musicological terms. Moreover, regardless of musical education, most people consciously or unconsciously seem to have some degree of musical understanding, making it a potentially useful tool for auditory interface design. Knowledge of musical building blocks, such as tempo, rhythm, melody, harmony and instrumentation, is therefore crucial to auditory interface designers.

Tempo, similarly to duration, is a relative parameter. Differences in tempo between two earcons can be difficult to recognise, unless they are subsequent. However, changes in tempo within an icon are easily perceived. A musical sequence that slows down or speeds up can for example indicate loading a battery or diminishing battery levels. Similarly, differences in dynamics and rhythm offer a great variability in auditory icons. Gaver (1997) suggests that the overall repetition rate of a rhythm may be used to convey a sense of activity. However, time limitations tend to constraint the possibilities of using rhythm, unless it indicates an ongoing event. Still, short rhythmic sequences can be differentiated on a basis of a few simple structures: iamb, dactyl, trochee, anapaest and amphibrach. Possibly, simple meters can also easily be distinguished.

When combining intervals over time melodies emerge. Short melodies are easily memorised and recognised, and constitute a great resource for auditory interface design. Additional musical dimensions can be added to an interface by using chords, in which tones of different pitches are played simultaneously. On a larger scale, harmonic movement emerges from chordal and melodic movement, although as with rhythm, the possibility to convey such complex information is often diminished by time limitations.

If the auditory feedback is allowed to span long durations of time, the possibilities of musical interfaces are virtually unlimited. One strategy is to compose musical signals that are based on a speech metaphor. By letting the sound of a musical message imitate the melody of a corresponding spoken message, earcons can be made informative without being too intrusive. Especially the way mothers speak to babies seems to contain some fundamental patterns, that are general to many cultures. A staccato is common in human warnings, such as

“bad, bad, bad!” (At the same time staccatos are common patterns in the warning sounds of birds.) Musical messages can also be expressive on a more abstract level. Music is often referred to as the “language of feelings” (Beardsley 1981). While the absence of a clear syntax is an important limit to this analogy, there are still similarities in the way people experience musical features. One example from the Western music culture is that minor keys are often associated with sad feelings, while major scales sound happier. More complex musical events can also be used. Modulating from the tonic key to the dominant and back can convey a feeling of departing and returning.

Peter Kivy (1984) argues that music alone might not be powerful enough to manipulate listeners’ feelings, but is well capable of evoking moods. However, associations can enhance the musical impact on feelings. Most people claim there are songs that make them feel in certain ways because the song relates strongly to an event in their lives. Such experiences are most often personal and hence not of any use for general interfaces. However, there are several examples of more general association effects. Music is often used in theatre, films and computer games to link events to feelings. Since most people are exposed through media to such musical clichés from early childhood, one could assume that significant conditioning effects exist. Similarly to the genre characteristics of common timbres mentioned above, melodies can also function as genre sounds. Gaver (1998) mentions that the theme from the emotion picture *Jaws* has become a well-known sound, which is linked to danger.

## 5 Conclusions

In current technology, there are many silent areas where auditory interfaces can be useful. When compared to the world around us, most interfaces are remarkably quiet. If interface designers would become more aware of the potentials of sound, auditory interfaces could be an important tool for enhancing human-machine interaction. When carefully designed, auditory interfaces can be effective and pleasant complements, or even substitutes, for visual interfaces. Since sound is fundamentally different to vision, it can convey information that vision cannot.

The many dimensions of sound make auditory interface design very flexible. The dimensions that are available are roughly: pitch, timbre, loudness, duration and direction. Non-speech auditory interfaces can either be based on everyday sounds or musical sounds. An interface can either be iconic, symbolic or metaphoric. Iconic representation generates the most efficient mapping, since it utilises the user’s

previous experiences. Realistic auditory icons are easily learned and remembered, but they can be difficult to develop. The interface developer should try to obtain a consistent strategy on how to convey meaning, which limits everyday sound as the foundation for auditory interfaces. Entirely symbolic auditory messages, on the other hand, are arbitrarily mapped to what they are to represent. Therefore, they are not as naturally integrated with visual interfaces, as are iconic mappings. Ideally, musical messages can be made metaphoric, or representational, meaning that they carry everyday sound attributes. Small changes in timbre and parameterising can create useful metaphors. Metaphoric sounds are in a way a compromise between a strictly iconic or symbolic approach to design.

There are several crucial issues when developing an auditory interface. The sounds used must be comprehensible and easy to remember. Every sound used must be easily distinguishable from one another as well as from ambient sound. The sounds should not be annoying, yet not too subtle to be noticed. If the user is to be frequently exposed to the auditory interface, aesthetic properties are fundamental. Therefore, excessive speech should be avoided, unless very specific information must be conveyed. Speech is also generally slow and might interfere when using communication devices such as phones. Since music is an aesthetic and complex structure for organising sound, it is a useful asset when designing long-lasting auditory interfaces.

An abstract sound interface combined with speech feedback can be a very powerful interactive aid. Certainly, there is a need for extensive tests and sociological studies in these areas, as interfaces always exist in a cultural context. One advantage of abstract auditory interfaces is that they are not language specific. However, even if music can be called "the language of feelings," one should not assume that it is universal. For example, Europeans and Africans might interpret the metaphors of ascending or descending intervals differently. Cultures such as the Chinese, where the language is sensitive to pitch variations, may be able to develop auditory icons on fundamentally different bases than Westerners. The phenomena of major and minor chords being associated with happy and sad moods are characteristic for Western culture only. Some musical traits however, such as pentatonic scales, are common in several different music cultures. Moreover, since Western music traditions today are widely spread, they could be a useful fundament for international auditory interfaces.

From an aesthetic perspective, interface users of different ages might have different musical preferences. As with many other aspects of human computer interaction, there seem to be lessons to learn from the world of computer games. Computer games generally emit more

sound than most other computer applications – and most users seem to enjoy it. Research on what makes good sound effects and music for computer games could provide many useful hints to other interface designers. There are many ways to control the amount of auditory feedback emitted. One way is to let sounds fade away over time and only reappear when the user asks for them or fails to perform the corresponding task. An alternative to this is a brief delay, so that the sound is only heard if the user hesitates. Naturally, there must always be an alternative to turn the sounds off.

There is much is yet to be explored about the learning process of interfaces. Association effects are often subjective, although common factors can be found, as is the case with genre sounds. In his experiments with earcons in 1998, Brewster found that test persons with previous musical training did not achieve better than others. Still, developing interfaces where musical education is beneficial is not downright discriminating. In fact, most existing interfaces do require some kind of background knowledge. Several depend on the user's literacy while others trust the user's capability of understanding visual symbols. As auditory interfaces prove to be useful, people ought to become more interested in requiring basic knowledge about sound and music.



# Reference List

Beardsley, Monroe C. 1981. "Understanding Music." In *On Criticizing Music: Five Philosophical Perspectives* ed Kingsley Price. Johns Hopkins.

Brewster, Stephen A. 1998, "Using Non Speech Sounds to Provide Navigation Cues." *ACM Transactions on Computer-Human Interaction* 5:3: 224 –259.

Gaver, William W. 1989. "The Sonic Finder" *Human-Computer Interaction* 4:1. Elsevier Science.

Gaver, William W. 1997. "Auditory interfaces." In *Handbook of Human-Computer Interaction* 2nd ed. 1003 –1041.

Kivy, Peter. 1984. "Representation and Expression in Music." In *Sound and Semblance*. Princeton University Press.

Norman, Donald A. 1990. *The Design of Everyday Things*. New York: Doubleday.

Wishart, Trevor. 1996. *On Sonic Art*. Revised ed. *Contemporary Music Studies Series* vol. 12. Ed. Simon Emmerson. Amsterdam: Harwood Academic.