Auditory Interfaces and Sonification

Michael A. Nees & Bruce N. Walker

Georgia Institute of Technology
Atlanta, GA, USA

Contact:

Bruce N. Walker
School of Psychology
Georgia Institute of Technology
654 Cherry Street
Atlanta, GA 30332-0170
USA

bruce.walker@psych.gatech.edu
Phone: +1-404 894-8265
Fax: +1-404-894-8905

Auditory Interfaces and Sonification

Auditory interfaces and sonification—information display by means of nonspeech audio (Kramer et al., 1999)—have been the subject of increasing interest in recent decades (for reviews, see Frysinger, 2005; Kramer et al., 1999). With the advent of ubiquitous digital technologies, high fidelity sound samples have become increasingly easy and inexpensive to produce and implement (Flowers, Buhman, & Turnage, 2005; Hereford & Winn, 1994). Perhaps more importantly, however, an increasing awareness of the shortcomings and limitations of traditional visual interfaces has spurred research on sound as a viable mode of information display. Nonspeech audio cues have been implemented to varying degrees in interface design, ranging from nonspeech audio as a complement or supplement to existing visual displays (e.g., Brewster, 1997; M. L. Brown, Newsome, & Glinert, 1989), to hybrid systems that integrate nonspeech audio with other audio technologies (e.g., screen readers, see Morley, Petrie, O'Neill, & McNally, 1999; Stockman, Hind, & Frauenberger, 2005). Attempts have even been made to develop interfaces (usually for the visually impaired) where feedback and interaction are driven primarily by sounds (e.g., Bonebright & Nees, 2007b; Edwards, 1989a, 1989b; Mynatt, 1997).

Despite the potential utility of sound in interface design, a recent survey of experts in HCI and usability (Frauenberger, Stockman, & Bourguet, 2007a) reported that only about 58% of respondents had designed with audio in any form. Nonspeech audio and sonification represent an important tool for universally accessible interface design, yet most interface designers consider speech audio first (and perhaps exclusively) when implementing audio in a system. Perhaps as a relic of the limited sound production capabilities of early personal computers (see Flowers et al., 2005), perceptions (and in some cases legitimate concerns) linger that sounds in interfaces are a minimally informative annoyance to the user.

We argue here that appropriately chosen and implemented nonspeech sounds can be a pleasant, informative, and integral part of interface design, and interfaces with nonspeech audio can promote adherence to at least 5 of the 7 principles of universal design (Connell et al., 1997; McGuire, Scott, & Shaw, 2006), including (1) equitable use; (2) flexibility in use; (3) simple and intuitive use; (4) perceptible information; and (5) tolerance for error**.**

The current chapter seeks to provide an introduction to nonspeech auditory information display and an overview of the relevant issues and critical decision points regarding the use of nonspeech audio in interfaces. Our discussion is guided by the theme that nonspeech auditory displays can universally enhance the human operator's experience with human-machine systems. Although we focus on the potential benefits of nonspeech audio, we refer the interested reader to other chapters in this volume (e.g., Chapter 38: Screen Readers, Chapter 40: Speech Interaction, and Chapter 51: Multimodality) for a complete discussion of the range of interface options available to the auditory or multimodal display engineer.

I. Appropriate Uses of Nonspeech Auditory Display

The best-practice use of nonspeech audio in interfaces requires a careful consideration of the types of users, tasks, and environments where the system will be implemented (for more detailed discussions, see Barrass, 1997; Kramer, 1994; Nees & Walker, 2007). To the extent that nonspeech audio is able to effectively convey the intended message, obvious accessibility benefits are incurred by certain types of system users (i.e., equitable use, see Connell et al., 1997; McGuire et al., 2006), particularly the 161 million people worldwide who are blind or visually impaired (Resnikoff et al., 2004). Screen readers (see Chapter 38 in this volume) have been quite effective at making text (and other verbal information) accessible for blind and visually impaired people across a wide variety of digital systems (Tobias, 2003). Other aspects of the interface (e.g., spatial, pictorial, or iconic information, etc.), however, cannot be easily represented with a simple text translation, and the inherent limitations introduced by a text-to-speech display system may introduce new navigation and usability difficulties, especially when the original materials (e.g., web pages, etc.) were not developed with a consideration of screen reader accessibility (see Mankoff, Fait, & Tran, 2005).

While accessibility for special populations has been one driving force in auditory display research, certain task dependencies and environmental conditions may render the affordances of nonspeech audio beneficial for all users of a system. For example, recent advances in technology have paradoxically expanded the realm of visual information display toward opposite extremes in physical size. Portable devices like cell phones, mp3 players, and even laptop computers continue the trend toward smaller physical dimensions, thereby leaving appreciably less space (or perhaps even no space) for a visual display to occupy (see, e.g., Brewster, 2002). Fixed work stations, on the other hand, have become characterized by multiple visual displays with increasingly large physical sizes, due in part to increases not only in the affordability of displays but also in the expanded computing power to support multiple concurrent displays. As a result, visually intensive workstations and other multitasking situations may overburden the visual modality (see Grudin, 2001). System limitations from both small and large visual displays are universally applicable and not unique to any particular type of user, and the inclusion of nonspeech audio in some interfaces can promote universal design principles such as flexibility in use and perceptible information (see Connell et al., 1997; McGuire et al., 2006).

In addition to these display-related interface design challenges, environmental conditions external to the system may impose further obstacles for the use of traditional, visual only displays. Line of sight with a visual display may be obscured (e.g., a firefighter in a smoke-filled room) or unstable (e.g., a jogger viewing an mp3 player's display). Other task dependencies may introduce additional demands on the human visual system that prevent the concurrent use of a visual display (e.g., when navigating or using mobile devices while walking, driving, or performing any other visually demanding task). Audition requires no physical or stable line-of sight with a display device (Kramer, 1994), which again allows for equitable use, flexibility in use, and perceptible information, and the inclusion of audio cues may even introduce more tolerance for error (see Connell et al., 1997; McGuire et al., 2006) into the system than visual displays alone.

Another notable property of the human auditory system is its sensitivity to the temporal aspects of sound (Bregman, 1990; Flowers, Buhman, & Turnage, 1997; Flowers & Hauer, 1995; Kramer, 1994; Kramer et al., 1999). In many instances, response times for auditory stimuli are faster than those for visual stimuli (Kramer, 1994; Spence & Driver, 1997). Furthermore, people can resolve subtle temporal dynamics in sounds more readily than in visual stimuli, thus the rendering of data into sound may manifest periodic or other temporal information that is not easily perceivable in visualizations (Flowers et al., 2005). Audition, then, may be the most appropriate modality for simple and intuitive (see Connell et al., 1997; McGuire et al., 2006) information display when data have complex patterns, express meaningful changes in time, or require immediate action.

II. A Brief Taxonomy of Nonspeech Audio and Sonification

While nonspeech audio has an important role to play in interface design, the specific types of nonspeech sounds that could be used to solve a given interface design challenge are numerous and diverse. Proposed categorical descriptions of nonspeech sounds generally have been arranged according to form (i.e., according to the parameters of the sound) or function (i.e., with respect to the role of the sound within a system) with some convergence between these approaches. A brief description of the types of nonspeech sounds used in interface design is offered here. We have chosen to organize our discussion roughly according to the functions of sounds in interfaces, but in reality the definitional boundaries for nonspeech audio sounds tend be vague and overlapping. For more discussion on taxonomic descriptions of nonspeech auditory displays, the interested reader is referred to Kramer (1994), Walker and Kramer (2004, 2006a, 2006b), and de Campo (2007), whose Sonification Design Map organized the relationships between nonspeech auditory displays along several quantitative continua.

Alarms, alerts, and warnings

Alarms, alerts, and warnings are generally brief, infrequent, unsubtle sounds designed to capture a person's attention. Traditionally alerts and warnings convey binary status information about an event's onset or offset (Edworthy & Hellier, 2006). For example, a doorbell informs a dwelling's occupants that someone is at the door (i.e., the alert indicates the onset of an event, the arrival of a visitor); this alert does not indicate who is outside, or what they might want. Alerts and warnings usually convey that immediate (or at least temporally proximal) action is required, and Haas and Edworthy (1996) found that higher frequency, rate, and intensity all contribute to more perceived urgency in an auditory alarm signal.

Object, item, and status indicators and auditory menus

Sounds such as earcons (e.g., Blattner, Sumikawa, & Greenberg, 1989; Bonebright & Nees, 2007a; Brewster, Wright, & Edwards, 1993; McGookin & Brewster, 2004), auditory icons (e.g., Bonebright & Nees, 2007a; Gaver, 1989; Keller & Stevens, 2004), and spearcons (Palladino & Walker, 2007; Walker, Nance, & Lindsay, 2006b) are examples of status and process indicators. Like alerts and warnings, these sounds tend to be brief, but they provide informative cues about the nature of the underlying action or event. These sounds are often used to facilitate tasks such as scrolling (Brewster, Wright, & Edward, 1994), pointing, clicking, and dragging with the mouse (Winberg & Hellstrom, 2003), or moving files, etc., in the interface. *Earcons* are abstract, artificial sounds that bear no ecological relationship to the represented process or event (e.g., beeps, chimes, abstract sound motives, etc., see Blattner et al., 1989). *Auditory icons* are more natural sounds that have some real world relationship with their referent process or event (Gaver, 1989), although the degree of ecological relatedness may vary (see Keller & Stevens, 2004). The abstract nature of earcons allows for flexibility in representation, as such abstract sounds can be assigned to most any object, item, or process in an interface. A trade-off exists, however, in that the user is required to learn the association between sounds and their referents; for large catalogs of abstract sounds, users may be unwilling or unable to learn the meaning of the sounds (Watson & Kidd, 1994). Research has shown that auditory icons are generally easier to learn and remember than earcons (Bonebright & Nees, 2007a; for a review, also see Edworthy & Hellier, 2006), but auditory icons are less flexible in that some objects, items, and processes have no inherent, ecological sound association (e.g., What sound should represent a "save" command?).

Recently, an alternative to earcons and auditory icons has emerged which may be able to ameliorate some of the flexibility-learnability trade-off in interface sounds. *Spearcons* use temporally compressed speech to represent objects, items, or processes with sound (Palladino & Walker, 2007; Walker et al., 2006b).[1] Spearcons have been shown to outperform both earcons

---

[1] Whether or not spearcons are recognized by listeners as speech may depend upon the listener's abilities and experience as well as the word or phrase that is accelerated. As the name implies, we view spearcons as a hybrid of speech and nonspeech auditory displays.

and auditory icons (Walker et al., 2006b) and may be especially useful in the design of flexible auditory menus (see Palladino & Walker, 2007), or for representing a large number of items.

Data representation and exploration

Rather than offering a brief indication of a transitory system state, auditory displays for data exploration use sound to represent information from an entire (usually quantitative) data set. *Auditory graphs* (for representative work, see L. M. Brown & Brewster, 2003; Flowers & Hauer, 1992, 1993, 1995; Nees & Walker, 2007; Smith & Walker, 2005) are typical examples of sonifications designed for data exploration purposes. Auditory graphs most commonly use changes in auditory frequency to correspond to changes in data values along the visual Y axis, while time corresponds to the visual X axis. Nees and Walker (2007) recently proposed a conceptual psychological model of auditory graph comprehension. They argued that the advantages of visual graphs, namely the emergence of otherwise unnoticed patterns and data features in plots of data, can be preserved in auditory representations of quantitative data. In much the same way as individual data points combine to form cohesive patterns in a visual graph, sequences of notes in auditory graphs are grouped according to Gestalt principles and can convey equivalent information (see Nees & Walker, in press).

Exploratory work has also examined auditory versions of numerous traditional display formats, including auditory scatterplots (e.g., Bonebright, Nees, Connerley, & McCain, 2001; Flowers et al., 1997), box-whisker plots (Flowers & Hauer, 1992; Peres & Lane, 2003; Peres & Lane, 2005), histograms (Flowers & Hauer, 1993), multidimensional data sets (see Hermann & Hunt, 2005), and tabular data (Stockman et al., 2005). These efforts have commonly relied on variations of the pitch-time display format described above, and the variety of displays that have been developed suggest an auditory analog for many visual graphical displays. *Audification*, for example, shifts the waveforms of periodic data into the audible range of frequencies for data exploration (e.g., seismographs, see Dombois, 2001). As an alternative to traditional visualizations, auditory displays of quantitative information may: (1) make data accessible for visually impaired students and scientists, thereby promoting collaborative efforts; (2) provide an immersive, multimodal, and more effective educational experience for all students of math and science; (3) allow for the detection of otherwise unnoticed patterns and anomalies in data; and (4) offer an equivalent, alternative mode of information display in circumstances where visual information display is inadequate (see, e.g., Kramer, 1994; Kramer et al., 1999; Nees & Walker, 2007). These advantages epitomize the spirit and principles of universal design (Connell et al., 1997; McGuire et al., 2006).

Spatial audio displays and the symbolic representation of spatial relationships in GUIs

A number of studies have confirmed that auditory signals can direct visual attention to a spatial location (e.g., Brock, Stroup, & Ballas, 2002; Eimer, 2001; McDonald, Teder-Salejarvi, & Hillyard, 2000; Mondor & Amirault, 1998, also see; Schmitt, Postma, & De Haan, 2000; Spence, McDonald, & Driver, 2004), and the spatial manipulations of audio have been shown to facilitate a three-dimensional visual search (Bolia, D'Angelo, & McKinley, 1999). Thus, spatial audio has been recognized as an important means of capturing, orienting, or guiding attention (Kramer, 1994). Current technology allows for the delivery of virtual spatial audio: a two point sound source (e.g., headphones) in conjunction with head-related-transfer-functions (HRTFs) can induce the perception that a sound originated from an external environmental source (Folds, 2006; Walker & Lindsay, 2005b; Wightman & Kistler, 1983, 1989).

In addition to orienting applications, virtual spatial audio cues have been successfully implemented as audio-only navigational aids, where the virtual spatial location of an audio beacon guides the user along a specified path to a destination.[2] Examples of this approach include the System for Wearable Audio Navigation (SWAN, Walker & Lindsay, 2005b; Wilson, Walker, Lindsay, Cambias, & Dellaert, 2007), and the Personal Guidance System (PGS, Golledge, Loomis, Klatzky, Flury, & Yang, 1991; Loomis, Golledge, & Klatzky, 1993; Loomis, Marston, Golledge, & Klatzky, 2005). The SWAN system generally employs spatialized nonspeech sounds, whereas the PGS has usually used spatialized speech.

Walker and Lindsay tested a number of different types of audio beacons, including pink noise, a sonar-like ping, and pure tones, and the broad-spectrum pink noise cue was found to be particularly effective for guiding navigation (Walker & Lindsay, 2005a, 2006). While a wealth of data support the feasibility of nonspeech audio as a navigation aid, it should be noted that performance outcomes for navigation were negatively impacted by the introduction of a (particularly difficult) concurrent speech discrimination secondary task (Walker & Lindsay, 2005a). The practical cost of these laboratory-induced performance decrements for the dual-task are unclear, and more research is needed to clarify how competing auditory signals may or may not result interference for navigation systems and indeed all auditory displays (see *Concurrent sounds* section below).

While spatial audio has been shown to effectively direct attention and guide navigation through physical space on a gross, or macro-level (e.g., from upwards of several inches), much research has been directed at the representation of spatial relationships with sound for smaller

---

[2] Virtual spatial audio cues seem particularly suited to indicate *where* an operator should look or move in physical space. We note, however, that attempts to map virtual audio spatial location to non-spatial data (e.g., using stereo panning and higher or lower virtual spatial elevation to represent quantities for conceptual dimensions, etc., see Roth, Kamel, Petrucci, & Pun, 2002) have been less successful, perhaps owing to systematic misperceptions of virtual elevation (see Folds, 2006).

physical spaces, such as the dimensions (i.e., the screen size) of traditional visual display. For example, lateralized audio (e.g., left-right stereo panning) has been used in conjunction with frequency cues (with higher frequency corresponding to higher spatial position) to provide auditory representations of the spatial relationships between objects on a computer screen (Winberg & Hellstrom, 2003). Other interfaces have used increasing pitch to represent movement from left to right and up and down on the screen (Edwards, 1989b), while yet other approaches have used combinations of pitch manipulations and the number of sounds presented to indicate position within a grid of rows and columns on a computer display (Bonebright & Nees, 2007b). Some of these projects have been targeted at visually impaired users; some have specifically targeted sighted users. Nevertheless, the approaches are inherently universal in that they promote alternative and flexible means of interaction with interfaces for all users.

Despite the insights gained from such studies, there remains no inherent, standard, or even clearly best way to use sound to convey the spatial relationships between objects in user interfaces. A major design dilemma, then, involves the extent to which audio interfaces should maintain the conventions of visual interfaces (Mynatt & Edwards, 1992), and indeed most attempts at auditory display seek to emulate or translate elements of visual interfaces to the auditory modality. While retrofitting visual interfaces with sound can offer some consistencies across modalities, the constraints of this approach may hinder the design of auditory interfaces, and native auditory interfaces would likely sound much different from interfaces designed with a relative visual counterpart in mind. While visual objects exist primarily in space, auditory stimuli occur in time. A more appropriate approach to auditory interface design, therefore, may require designers to focus more strictly on auditory capabilities. Such interfaces may present the items and objects of the interface in a fast, linear fashion over *time* (see, for example, Eiriksdottir, Nees, Lindsay, & Stanley, 2006) rather than attempting to provide auditory versions of the spatial relationships found in visual interfaces. This approach can often lead to the deployment of enhanced auditory menus with a mix of speech and non-speech components. Such advanced interfaces are relatively novel (compared to simpler text-to-speech menus). Ongoing research in advanced auditory menu-based interfaces looks promising, and will generally provide better interfaces for all users (see Yalla & Walker, 2007).

Soundscapes and background auditory displays

Many continuous auditory stimuli can be allowed to fade to the extreme periphery of conscious awareness, yet meaningful changes in such on-going sounds are still noticed (Kramer, 1994). Designers have taken advantage of this auditory capability with *soundscapes*—ambient, continous sonifications—to facilitate a human operator's awareness of dynamic scenarios (e.g., a bottling plant, Gaver, Smith, & O'Shea, 1991; financial data, B. S. Mauney & Walker, 2004; a crystal factory, Walker & Kramer, 2005). Soundscapes often have been designed to mimic natural, ongoing auditory stimuli (e.g., a thunderstorm with rain), and parameters of the soundscape are mapped to particular variables in a multidimensional data set (e.g., B. S. Mauney & Walker, 2004). While the listener may not necessarily act upon every change in the soundscape, the display allows for on-going monitoring and awareness of a changing situation.

Arts and entertainment

Researchers and musicians have long recognized the potentially unique aesthetic or entertainment value of data-driven (i.e., sonified) music[3] (see, for example, Quinn, 2001), and the International Conference on Auditory Display has regularly featured a concert performance (e.g., , "Global music - The world by ear," 2006; "Listening to the mind listening: Concert of

---

[3] Also see http://www.tomdukich.com/weather%20songs.html

sonifications at the Sydney Opera House," 2004). A recent push in research, however, has taken the notion of sonification as entertainment a step further by advocating for enhanced and accessible exhibitions (e.g., museums, aquaria, zoos, etc.). People with disabilities, particularly the visually impaired, have been shut out of many of the educational and entertainment ("edutainment") experiences offered at traditional exhibitions. While virtual, on-line accessible museums are one possible solution to the problem (see Anable & Alonzo, 2001), a remote virtual experience lacks many important aspects (including the novelty and excitement) of a live visit to the actual sites of educational and culturally meaningful exhibitions. While recommendations for real museum accessibility are available (Salmen, 1998), the audio component of accessibility has primarily involved text-to-speech conversions of plaques and verbal materials—a practice which does not capture the most interesting aspects of dynamic exhibitions.

Walker and colleagues (Walker, Godfrey, Orlosky, Bruce, & Sanford, 2006a; Walker, Kim, & Pendse, 2007) have recently begun developing a system for sonifying the real-time dynamics of an aquarium. The movements of the fish are tracked (e.g., with computer vision) and translated to continuous, non-speech (and often musical) auditory representations. The result is a soundscape whereby categorical information about the types of fish can, for example, be represented by instruments of different timbre, while movements of the fish can be conveyed by other dimensions of sound such as pitch, tempo, loudness, or spatial location. Similar innovative approaches may enhance the experience of both static and dynamic exhibitions for all users, as supplementary audio may provide for a more immersive environment in museums, zoos, and aquaria where line-of-sight contact with the exhibit may be obscured by crowds or by perceptual or mobility impairments.

Another important development in accessible entertainment has been an increased interest in auditory games (also see Chapter 56: Developing Universally Accessible Games, in this volume). Audio-only interfaces have been developed for traditionally visual games such as the Towers of Hanoi (Winberg & Hellstrom, 2001) and Tic-Tac-Toe (Targett & Fernstrom, 2003). More elaborate attempts at audio-only gaming have also begun to appear, including an auditory role-playing game based on the Beowulf story (Liljedahl, Papworth, & Lindberg, 2007). Liljedahl et al. argue that audio-only gaming offers players the opportunity to construct rich, unconstrained internal images of the game's landscape from the suggestive nature of the sounds. Interestingly, a recent prototype for an audio-only computer soccer game may actually be able to offer constructive insights for both blind and sighted players on the real soccer field (Stockman, Rajgor, Metatla, & Harrar, 2007).

III. Design Considerations for Auditory Interfaces

Theoretical accounts of human interactions with sonification and other nonspeech auditory display design have been slow to develop, in part due to the highly interdisciplinary nature of the field (Nees & Walker, 2007). Recently, however, a number of authors have taken steps toward elaborating sonification theory and organizing the extant knowledge base, including de Campo's Sonification Design Space Map (de Campo, 2007), Frauenberger, Stockman, and Bourguet's audio design survey (2007a) and framework (2007), and Nees and Walker's model of auditory graph comprehension (2007). Despite these recent advances in the field, concrete and specific sonification design guidelines that are grounded in literature and theory are still not generally available. While researchers have described guidelines for nonspeech auditory displays (L. M. Brown, Brewster, Ramloll, Burton, & Riedel, 2003; Edworthy & Hellier, 2006; Flowers, 2005; Hereford & Winn, 1994; Watson & Kidd, 1994), these attempts have generally provided advice for particular instantiations of auditory displays as opposed to generalized recommendations or comprehensive descriptions for the entire scope of nonspeech audio. Furthermore, in at least one case it has been shown that adherence to published standards for auditory displays *did not* even ensure the identifiability of sounds (see Lacherez, Seah, & Sanderson, 2007). Rather than articulating what would necessarily be an incomplete list of rules or guidelines here, we instead offer a broader discussion of the critical issues for implementing nonspeech audio in interface design. We suggest that a careful consideration of these topics will help to ensure the appropriate deployment of sound in a system and offer a universally accessible and enhanced interface experience for all populations of users.

Detectability and discriminability

An auditory display is useless if the listener cannot hear the sounds in the system's environment of operation. Research in psychoacoustics has provided ample descriptions of minimum thresholds for detection of sounds along a number of relevant auditory dimensions (e.g., Hartmann, 1997), while masking theories have made valuable predictions about the human listener's ability to hear a sound signal against noise (for a discussion, see Watson & Kidd, 1994). The highly controlled testing conditions for such stimuli, however, can be drastically different from the environments where auditory displays will actually be used by listeners. Accordingly, ecologically plausible testing conditions for applications of auditory displays have been recommended (Brewster, 2002; also see Walker & Kramer, 2004; Watson & Kidd, 1994). Another concern is central or informational masking, whereby sounds are masked at higher levels beyond the cochlea in the auditory system. This variety of masking is not well understood, nor can it readily be predicted by extant models of the acoustic periphery (see Durlach et al., 2003). While the requirement of detectability for auditory information may seem straightforward, the interface designer may encounter problems if simple detection is not given due consideration during the design process.

Given that a sound can be heard by the human listener in the system's environment of operation, a second basic consideration is the discriminability of sounds with distinct meanings in the interface. Like detection, researchers have studied the discriminability of sounds along a wealth of dimensions such as pitch (e.g., Stevens, Volkmann, & Newman, 1937; Turnbull, 1944), loudness (e.g., Stevens, 1936), tempo (e.g., Boltz, 1998), and duration (e.g., Jeon & Fricke, 1997), to name but a few. Again, however, the stimuli and controlled conditions for data collection in such studies may not precisely translate to the real world scenarios where auditory interfaces will be used, and the designer is cautioned to proceed with an awareness of both the psychoacoustic discriminability of manipulated dimensions of sounds as well as the further additional constraints imposed by the tasks and environments for which the system is designed. Two sounds that carry different pieces of information must be distinguished to ensure that the operator will perceive the intended message.

Annoyance

The potential for sounds to annoy the user is a concern for auditory interface design (Frauenberger, Stockman, & Bourguet, 2007b; Kramer, 1994). Edworthy (1998) described the independent nature of sound aesthetics and performance outcomes. Sounds that annoy the user may be ignored or turned off, even when the presence of auditory cues enhances user performance with the system. Likewise, sounds may enhance the aesthetic experience of an interface without improving performance with the system. Some have suggested that musical nonspeech sounds (e.g., sounds from the MIDI instrument base) with their richer harmonic and acoustic features are easier to perceive than pure tones and simple waveform sounds (L. M.

Brown et al., 2003; Childs, 2005; Ramloll, Brewster, Yu, & Riedel, 2001). Simply using musical sounds, however, will not guarantee a pleasant experience of the auditory interface for all users, tasks, and environments. Bonebright and Nees (2007b) recently found that four different types of earcons (including both pitched musical instruments and pure-tone based variations) as well as a speech condition all led to auditory displays that were rated as "neutral" to somewhat "annoying" in the context of the study task, which was a dual-task listening and orienting paradigm. Another study found that high-pitched interface sounds can be particularly annoying (Bonebright & Nees, 2007a). This makes it clear that developing an auditory interface is, in all regards, a design task, with all the inherent difficulties associated with design. It is encouraging, however, that other research has shown that that users can be very satisfied with abstract, nonspeech sounds similar to those used by Bonebright and Nees (e.g., Morley et al., 1999).

In general, very little research has addressed the role of aesthetics in auditory display design and many questions remain regarding how to make aesthetically pleasing interface sounds. It remains advisable to pilot sounds with a representative sample of the target user group in order to eliminate particularly annoying and displeasing sounds, unless such sounds are invoked with a specific intent (e.g., as an alarm tied to a critical, rare event, etc.). Another possible solution involves customizability, where users are given a choice of instruments or sound types, all of which can convey equivalent information. Regardless of the approach, evaluation of aesthetics needs to be longitudinal, since preferences can evolve, and acceptance can increase or decrease as the user becomes more familiar with the interface.

Mappings, scalings, polarities

Mapping refers to the dimension of sound that is employed to vary with and thus represent changes in data. In general, groups of listeners have shown some concurrence about which aspects of sound are good for portraying certain conceptual dimensions of data. Nees and Walker (2007) give a detailed discussion and justification of the convention of mapping pitch to Y-axis spatial location in auditory graphs, and pitch generally offers a robust mapping dimension for quantities (also see L. M. Brown et al., 2003; Flowers, 2005). Some sound dimensions (e.g., loudness) are often not very effective representations of data for both perceptual and practical reasons (Neuhoff, Kramer, & Wayand, 2002; Walker & Kramer, 2004). Walker has attempted to determine the appropriate acoustic dimension for a given type of data by examining mappings between numerous conceptual data dimensions (e.g., temperature, pressure, danger) and three acoustic dimensions (pitch, tempo, and spectral brightness; Walker, 2002, 2007). Pitch, for example, generally maps well to changes in temperature, but tempo is not particularly effective for this conceptual dimension. Future research should extend and expand upon this approach to guide interface designers toward best-practice mapping choices. Currently, designers should be warned that not all acoustic mappings are equally effective for representing a given conceptual data dimension, and best-practice design decisions for interfaces will arise from an awareness of empirical data and usability pilot testing. As auditory display design requires explicit decisions regarding mapping, a variety of sources should be consulted to attain an awareness of the varieties of mappings available for nonspeech auditory display designers. (e.g., Bonebright et al., 2001; L. M. Brown et al., 2003; Edworthy, Hellier, Aldrich, & Loxley, 2004; Flowers, 2005; Neuhoff et al., 2002; Walker, 2002, 2007). Redundant or dual mappings (i.e., mapping more than one acoustic dimension to changes in data) may further facilitate comprehension of the display (Bonebright & Nees, 2007a; Kramer, 1994).

Following the selection of an acoustic mapping for data, the polarity of the data-to-display relationship must be considered. Increases in a given acoustic dimension (e.g., pitch, tempo, etc.) are most often mapped to increases in the data represented (a positive mapping polarity, Walker, 2002, 2007), but listeners agree that some conceptual data dimensions are better represented with a negative polarity mapping. For example, listeners might agree that increasing pitch suggests increasing temperature, yet the same group of listeners may feel that decreasing pitch offers a more intuitive representation of increasing size. Walker and Lane (2001) showed that some polarity mappings were reversed for visually impaired as compared to sighted listeners. While positive polarities may generally capture listener intuitions (L. M. Brown et al., 2003), interface designers should be mindful of user populations and conceptual data dimensions for which this convention is violated. Walker (2002, 2007) provided data for the preferred polarities for many conceptual data dimensions, and usability testing is advisable when evidence regarding a specific polarity relationship is not available.

Along with polarity, the auditory display designer must also consider the amount of change in an acoustic dimension that will be used to represent a unit of change in the data. Magnitude estimation has been employed to describe the intuitive slopes for scaling frequency to a number of conceptual data dimensions (Walker, 2002, 2007), and the conceptual data dimension being represented impacts the choice of scaling factor in the display. For example, equal quantitative changes (e.g., a one unit increase) in different conceptual data dimensions (e.g., temperature and size) are not necessarily best represented by the same change in the acoustic display dimension. A match between the listener's preferred or intuitive internal scaling function and the display's scaling function may facilitate comprehension of the information presented, particularly when judgments of absolute or exact values are required. Where feasible, scaling factors should be chosen to match the intuitive user preferences for representing change in a given conceptual dimension (for a number of empirically determined scaling slopes, see Walker, 2002, 2007). Brown et al. (2003) have further suggested minimum (MIDI note 35, ~61.7 Hz) and maximum (MIDI note 100, ~2637 Hz) scaling anchors. Again, the interface designer is encouraged to consult available empirical data as guidance, but ultimately empirical findings, design experience and expertise, and usability pilot testing will converge to determine the best-practice for a given application.

Interactivity

Interfaces for different scenarios may vary considerably in the degree to which interactivity is allowed or encouraged. Some auditory interfaces, such as alarms, may simply be activated by a particular system condition and occur without any opportunity for the user to actively adjust or manipulate the display; non-interactive sounds in interfaces have been referred to as "tour based" (Franklin & Roberts, 2004) or "concert mode" (Walker & Kramer, 1996). Other auditory components of an interface may allow for the particular sound message to be replayed, etc., while displays at the extreme end of the interactivity spectrum may allow for elaborate user control of and immersion in the display, including pausing, scanning, scrubbing, skipping backward and forward, and zooming in and out of display dimensions. Such interactivity, called "query based" (Franklin & Roberts, 2004) or "conversation mode" (Walker & Kramer, 1996), may be especially helpful for tasks involving data exploration and analysis (see L. M. Brown, Brewster, & Riedel, 2002). For a type of auditory displays called *model-based sonifications*, user control is imperative and drives the presentation of sounds in an entirely active data exploration process (see Hermann & Hunt, 2005). The inclusion of interactive control over the auditory components of a system warrants a consideration of the role of audio in the system and the extent to which such features aid the user in the task at hand versus the cost and potential negative effects of building interactive control into the interface. Of course, the controls that enable this interactivity should also be designed to support universal access (see, for example, Chapter 39: Virtual Mouse and Keyboard for Text Entry and Chapter 43: Haptic Interaction in this volume).

Individual Differences

Important individual differences may influence the interpretation of auditory displays such that different users may interpret the same sounds to have different meanings. If a technology strives toward universal accessibility, the interface designer should be aware of the range and variety of individual differences that must be accommodated (e.g., Meyer & Rose, 2000). For universal design, then, individual differences represent not only a crucially important design challenge, but also an opportunity to meet the needs of diverse populations of users.

Individual difference variables that may be relevant to the interpretation of auditory displays include cognitive abilities (e.g., memory and attention), musical ability, listening skills, learning styles, and perceptual abilities. Walker and Lane (2001) found differences between groups of visually impaired and sighted listeners in magnitude estimation tasks. As mentioned above, this study indicated that in some situations visually impaired and sighted listeners intuit the same polarities for data-to-display mappings, but in other cases different polarities result. Sighted individuals, for example, preferred a positive polarity when mapping frequency to the conceptual dimension “number of dollars”, whereas visually impaired individuals preferred a negative polarity. Auditory interface designers must empirically examine and anticipate these potential conflicting intuitions across user groups and take caution against unknowingly creating a display that is biased against universal access.

Researchers have further suggested that the transient nature of auditory displays may impose inordinate burdens on memory (Frauenberger et al., 2007b; Morley et al., 1999), a concern that warrants a consideration of the impact of cognitive abilities (e.g., memory, attention, etc.) on auditory display performance. Walker and Mauney (2004) studied the impact of individual differences in cognitive abilities on auditory magnitude estimation tasks. They found some evidence that cognitive abilities affected the interpretation of auditory displays. Listeners with better scores on working memory capacity (WMC) and nonverbal reasoning measures performed better on the magnitude estimation task than those listeners who had lowers scores of WMC and spatial reasoning, however, the scaling slope of the data-to-display mappings did not seem to be affected by cognitive abilities, musical experience, or demographic variables (Walker & Mauney, 2004). Mauney (2006) investigated cognitive abilities and musical experience as predictors of frequency and tempo discrimination. Participants completed the Operation Span (O-span) task as a measure of working memory capacity and the Raven's Progressive Matrices task as a measure of nonverbal reasoning. The results of the regression analyses show that performance on the Raven's and O-span tests seemed to predict some, but not all, tested frequency and tempo discrimination thresholds, with better cognitive abilities associated with lower thresholds. As this pattern of results suggested, the role of cognitive abilities in the comprehension of auditory displays is not well understood, although there is reason to believe that further research will yield stable relationships between certain cognitive abilities and performance with auditory stimuli. The generally transient nature of auditory displays may impose memory demands that could exacerbate individual differences in cognitive variables, so good auditory interface design will require the intuitive use of audio that does not require memorization of large catalogs of sounds.

Researchers have long predicted that the special training and listening abilities of musicians would translate to superior performance with auditory displays as compared to nonmusicians, and a few studies have found such a relationship (e.g., Lacherez et al., 2007; Neuhoff et al., 2002; Sandor & Lane, 2003). In general, however, many researchers have reported weak to non-existent relationship between musical experience and performance with auditory displays (see Bonebright et al., 2001; Nees & Walker, in press; Walker, 2002; Watson & Kidd, 1994**.** Watson and Kidd (1994) suggested that the comprehension of auditory displays may simply require perceptual acuity (as opposed to musical ability per se), which is likely a variable that is distributed homogeneously across musicians and nonmusicians. Furthermore, while nonmusicans may not be formally trained in music theory, most adult listeners have at least acquired a wealth of *implicit* knowledge about the rules, structures, and relationships between sounds in music (Bigand, 1993). This implicit knowledge may be enough to perform tasks with auditory interfaces, which generally require no responses related to explicit musical knowledge. Finally, no brief, valid tools exist for measuring musical ability, and the use of surrogate measures (e.g., self-reported years of musical experience) may not be capturing enough of the variance in actual musical ability to detect meaningful relationships (Nees & Walker, 2007).

Training and skill acquisition

While accessibility often implies that a system should be intuitive and easily understood even by novice users, novel interfaces such as nonspeech auditory displays may require at least some minimal explanation or instruction for the user. Watson and Kidd (1994) accurately pointed out that many people will be unwilling to commit to extensive training in the meaning of sounds in an interface, yet brief training (i.e., under 30 min) has been shown to positively impact performance with auditory displays. Smith and Walker (2005) showed that brief training for a point estimation task resulted in better performance than no training. Walker and Nees (2005b) also demonstrated that a brief training period reduced performance error by 50% on a point estimation sonification task. Although to date, little attention has been paid to the issue of training sonification users, recent and ongoing work is examining exactly what types of training methods are most effective for different classes of sonifications (e.g., Walker & Nees, 2005a). Studies that have explicitly analyzed performance data over time (i.e., across trials or blocks of trials) have suggested that performance improves with more experience with the novel displays (Bonebright & Nees, 2007b; Nees & Walker, in press; Walker & Lindsay, 2006), but the upper limits of performance with auditory displays remain unknown (Nees & Walker, 2007; Walker & Nees, 2005a). Longitudinal studies of skill acquisition with auditory interfaces are needed.

Concurrent sounds

Numerous studies have shown that the discriminability and identifiability of sounds decreases as the number of concurrently presented sounds increases (Bonebright et al., 2001; Ericson, Brungart, & Simpson, 2003; Lacherez et al., 2007; McGookin & Brewster, 2004; Walker & Lindsay, 2005a). Theory and research alike suggest, however, that such problems can be somewhat ameliorated to the extent that acoustic cues allow for concurrently presented sounds to be parsed into separate streams (Bregman, 1990). To this end, researchers have suggested that spatial separation of different data (e.g., presenting different data series to left and right headphone channels, see Bonebright et al., 2001; L. M. Brown et al., 2003), the use of distinct timbres for different data series (Bonebright et al., 2001; McGookin & Brewster, 2004), and staggering the onsets of concurrent messages (McGookin & Brewster, 2004) may all facilitate the segregation of concurrent audio information. While pitch is also an effective cue for parsing concurrent auditory streams, pitch is often used to represent dynamic, non-categorical information in auditory displays and often may not be an appropriate dimension for promoting the separation of different data series. We should further note that the concurrent presentation of distinct channels of auditory information probably has a limit, beyond which the distinct streams of information will become impractical to parse and perceive (Flowers, 2005). This theoretical limit is likely dependent upon not only the number of concurrently presented sounds, but also upon their qualitative characteristics. Bonebright and Nees (2007b), for example, found little to no interference for the comprehension of speech passages in the presence of a concurrent orienting task with earcons. Care should be taken when interfaces or environmental circumstances allow for overlapping sounds, as more research is needed to clarify the limits of perception for simultaneous auditory input.

Delivery of audio: Hardware

A growing majority of digital devices come equipped with high-fidelity sound production capabilities off-the-shelf. The auditory component of many interfaces may require little or no modification to hardware, but rather a design philosophy that takes better advantage of the existing capability to improve system accessibility with audio. The hardware considerations for the delivery of an audio interface may vary across different use scenarios, however, and Walker and Lindsay (2005b), described a number of the challenges encountered when designing their system for wearable audio navigation (SWAN). The SWAN project encountered logistical constraints beyond the auditory interface itself, including technical limitations such as unreliability in sensors (e.g., the fallibility of GPS and other technologies that attempt to precisely determine a mobile user’s location) as well as practical limitations in battery power, size, and durability of a wearable, mobile computer (also see Chapter 11: Hand-held Devices and Mobile Phones and Chapter 12: Wearable Computing in this volume).

Similarly, many attempts at auditory interfaces have been coupled with custom input devices (e.g., Morley et al., 1999; Winberg & Hellstrom, 2003), but such improvisations may not be necessary to the success of an auditory interface. While novel or emerging hardware technologies may eventually transform many of the ways in which people interact with a system (see Chapter 39: Virtual Mouse and Keyboard for Text Entry and Chapter 43: Haptic Interaction, in this volume), we believe that the existing, off-the-shelf capabilities of most hardware already allows for the implementation of auditory interfaces that could offer enhanced accessibility for all users.

For delivering sound, audio-capable systems have traditionally relied upon speakers or headphones, both of which are inexpensive options for producing audio of sufficient fidelity for most applications of auditory displays. Privacy and the potentially intrusive nature of delivering sounds through speakers are inter-related, basic concerns. When used in the presence of other people, speakers not only may compromise a user's privacy, but also can interfere with the activities of those nearby or cause annoyance. Headphones may circumvent these problems, but having one's ear's covered by headphones introduces new difficulties for interacting with and maintaining awareness of one's surroundings. Blind users, for instance, gather a majority of their environmental information through from sound, and they will generally be unwilling to cover their ears, even to use a potentially beneficial system.

One potential solution that is actively being researched is bonephones—bone-conduction headphones. Small transducers sit on the mastoid behind the ear and vibrate the skull, effectively stimulating the cochlea directly and by-passing the outer and middle ear. The ears remain uncovered, but the delivery of private audio messages is still possible. Within minimal equalization, bonephones have been shown to have similar psychoacoustic signatures as headphones with regard to thresholds (Walker & Stanley, 2005), and early research suggests that virtual spatialized audio is possible with bonephones (Stanley & Walker, 2006). The devices, however, are currently not widely available to consumers, and more research is needed to clarify the potential role for interference between audio delivered via bonephones and concurrent stimulation from environmental sound sources.

Delivery of audio: Software

As described throughout this chapter, most digital devices have off-the-shelf hardware and software capabilities for sound production, and the success of auditory interfaces will primarily be a function of empirically-based design philosophies that embrace the use of sound. No standard add-on software packages exist for the general production of custom nonspeech audio for use in interfaces. Many laboratories involved in research on auditory displays, however, have developed purpose-specific sonification software packages that are often available as free, open-source downloads. Applications for representing data with sound include NASA's Mathtraxx[4], the Oregon State University Science Access Project's Accessible Graphing Calculator[5], and the Georgia Tech Sonification Lab's Sonification Sandbox[6] (see Davison & Walker, 2007; Walker & Cothran, 2003) and Auditory Abacus[7] (Walker, Lindsay, & Godfrey, 2004). Stockman and colleagues (Stockman et al., 2005) are working on a prototype for a software package that works with Microsoft Excel and CSound to allow for sonifying cells of spreadsheets, while Hetzler and Tardiff (2006) have also developed an Excel plug-in for data sonification. Cook (2007) recently described a number of software development projects aimed at analyzing and synthesizing environmental sounds. Other resources of interest for the auditory

---

[4] http://prime.jsc.nasa.gov/mathtrax/

[5] http://dots.physics.orst.edu/calculator.html

[6] http://sonify.psych.gatech.edu/~ben/sandbox/index.html

[7] http://sonify.psych.gatech.edu/research/audio_abacus/index.html

interface designer include the website of the International Community for Auditory Display[8] as well as the AUDITORY electronic mail list[9], both of which offer access to experts with years of collective experience in implementing sounds for research and application.

Conclusions

Sonification and auditory interfaces are uniquely suited to enhance and improve the universal accessibility of a system for a number of users, tasks, and environments. The thoughtful and informed addition of nonspeech audio to an interface, especially as one important element of a holistic approach to universal design, can enhance and improve the accessibility and usability of a system. Nonspeech audio is uniquely suited to convey particular types of information and to ameliorate some of the limitations imposed by traditional visual interfaces. For the visually impaired, computers and other digital technologies have dramatically impacted and will continue to improve access to education, employment, and an overall higher quality of life (Gerber, 2003; Tobias, 2003), and nonspeech auditory displays can fill gaps in accessibility related to alerting or warning functions, status or process updates, ongoing monitoring tasks, and even data exploration. The relevance of nonspeech audio to interface design extends well beyond affordances for the visually impaired (e.g., Griffith, 1990). The benefits of universally usable interfaces should extend system capabilities for all users during visually intensive tasks or in environments where vision is not the ideal modality for information display. Auditory interfaces and sonification can be major contributors to compliance with at least 5 of the 7 principles of universal design (Connell et al., 1997; McGuire et al., 2006), including (1) equitable use; (2) flexibility in use; (3) simple and intuitive use; (4) perceptible information; and (5) tolerance for error.

---

[8] www.icad.org

[9] www.auditory.org

Flowers (2005) asked whether sound should be a standard component of desktop interfaces. We believe that nonspeech sound is an under-used and under-investigated tool for the development of universally accessible interfaces. We further suggest that audio can be implemented immediately and cheaply in most existing interfaces, with little or no modifications to existing software and hardware. Ultimately, the potential of sound to benefit all users of an interface will only be unlocked when researchers commit to explore the best-practice role of sound in interfaces and when designers actively implement audio in interfaces.

References

Anable, S., & Alonzo, A. (2001). Accessibility techniques for museum web sites. *Proceedings of Museums and the Web (MW 2001)*, Seattle, WA.

Barrass, S. (1997). Auditory information design. Unpublished Dissertation, Australian National University.

Bigand, E. (1993). Contributions of music to research on human auditory cognition. In S. McAdams & E. Bigand (Eds.), *Thinking in Sound: The Cognitive Psychology of Human Audition* (pp. 231-377). New York: Oxford University Press.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human-Computer Interaction*, 4, 11-44.

Bolia, R. S., D'Angelo, W. R., & McKinley, R. L. (1999). Aurally aided visual search in three-dimensional space. *Human Factors*, 41(4), 664-669.

Boltz, M. G. (1998). Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure *Perception & Psychophysics*, 60(8), 1357-1373.

Bonebright, T. L., & Nees, M. A. (2007a). Memory for auditory icons and earcons with localization cues. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 419-422), Montreal, Canada.

Bonebright, T. L., & Nees, M. A. (2007b). Using an orienting auditory display during a concurrent listening task. *Manuscript submitted for publication.*

Bonebright, T. L., Nees, M. A., Connerley, T. T., & McCain, G. R. (2001). Testing the effectiveness of sonified graphs for education: A programmatic research project. *Proceedings of the International Conference on Auditory Display (ICAD2001)* (pp. 62-66), Espoo, Finland.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press.

Brewster, S. (1997). Using non-speech sound to overcome information overload. *Displays*, 17, 179-189.

Brewster, S. (2002). Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing*, 6(3), 188-205.

Brewster, S., Wright, P. C., & Edward, D. E. (1994). The design and evaluation of an auditory enhanced scrollbar. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 94)* (pp. 173-179), Boston, MA.

Brewster, S., Wright, P. C., & Edwards, A. D. N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 93)* (pp. 222-227), Amsterdam.

Brock, D., Stroup, J. L., & Ballas, J. A. (2002). Using an auditory display to manage attention in a dual task, multiscreen environment. *Proceedings of the International Conference on Auditory Display (ICAD2002)*, Kyoto, Japan.

Brown, L. M., & Brewster, S. (2003). Drawing by ear: Interpreting sonified line graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 152-156), Boston, MA.

Brown, L. M., Brewster, S., Ramloll, R., Burton, M., & Riedel, B. (2003). Design guidelines for audio presentation of graphs and tables. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 284-287), Boston, MA.

Brown, L. M., Brewster, S., & Riedel, B. (2002). Browsing modes for exploring sonified line graphs. *Proceedings of the 16th Britich HCI Conference*, London, UK.

Brown, M. L., Newsome, S. L., & Glinert, E. P. (1989). An experiment into the use of auditory cues to reduce visual workload. *Proceedings of the ACM CHI 89 Human Factors in Computing Systems Conference (CHI 89)* (pp. 339-346).

Childs, E. (2005). Auditory graphs of real-time data. *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland.

Connell, B. R., Jones, M., Mace, R., Mueller, J., Mullick, A., Ostroff, E., et al. (1997). The Principles of Universal Design, Version 2.0. NC State University, Raleigh, NC: The Center for Universal Design.

Cook, P. R. (2007). Din of an "iquity": Analysis and synthesis of environmental sounds. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 167-172), Montreal, Canada.

Davison, B. K., & Walker, B. N. (2007). Sonification Sandbox reconstruction: Software standard for auditory graphs. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 509-512), Montreal, Canada.

de Campo, A. (2007). Toward a data sonification design space map. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 342-347), Montreal, Canada.

Dombois, F. (2001). Using audification in planetary seismology. *Proceedings of the International Conference on Auditory Display (ICAD2001)* (pp. 227-230), Espoo, Finland.

Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, 113(6), 2984-2987.

Edwards, A. D. N. (1989a). Modelling blind users' interactions with an auditory computer interface. *International Journal of Man-Machine Studies*, 30, 575-589.

Edwards, A. D. N. (1989b). Soundtrack: An auditory interface for blind users. *Human-Computer Interaction*, 4, 45-66.

Edworthy, J. (1998). Does sound help us to work better with machines? A commentary on Rautenberg's paper 'About the importance of auditory alarms during the operation of a plant simulator'. *Interacting with Computers*, 10, 401-409.

Edworthy, J., & Hellier, E. (2006). Complex nonverbal auditory signals and speech warnings. In M. S. Wogalter (Ed.), *Handbook of Warnings* (pp. 199-220). Mahwah, NJ: Lawrence Erlbaum.

Edworthy, J., Hellier, E. J., Aldrich, K., & Loxley, S. (2004). Designing trend-monitoring sounds for helicopters: Methodological issues and an application. *Journal of Experimental Psychology: Applied*, 10(4), 203-218.

Eimer, M. (2001). Crossmodal links in spatial attention between vision, audition, and touch: evidence from event-related brain potentials. *Neuropsychologia*, 39, 1292-1303.

Eiriksdottir, E., Nees, M. A., Lindsay, J., & Stanley, R. (2006). User preferences for auditory device-driven menu navigation. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 2076-2078), San Francisco, CA.

Ericson, M. A., Brungart, D. S., & Simpson, B. D. (2003). Factors that influence intelligibility in multitalker speech displays. *International Journal of Aviation Psychology*, 14(3), 313-334.

Flowers, J. H. (2005). Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions. *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland.

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (1997). Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors*, 39(3), 341-351.

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (2005). Data sonification from the desktop: Should sound be part of standard data analysis software? *ACM Transactions on Applied Perception*, 2(4), 467-472.

Flowers, J. H., & Hauer, T. A. (1992). The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuocentric? *Behavior Research Methods, Instruments & Computers*, 24(2), 258-264.

Flowers, J. H., & Hauer, T. A. (1993). "Sound" alternatives to visual graphics for exploratory data analysis. *Behavior Research Methods, Instruments & Computers*, 25(2), 242-249.

Flowers, J. H., & Hauer, T. A. (1995). Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors*, 37(3), 553-569.

Folds, D. J. (2006). The elevation illusion in virtual audio. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1576-1580), San Francisco, CA.

Franklin, K. M., & Roberts, J. C. (2004). A path based model for sonification. *Proceedings of the Eighth International Conference on Information Visualization (IV '04)* (pp. 865-870).

Frauenberger, C., Stockman, T., & Bourguet, M.-L. (2007a). A Survey on Common Practice in Designing Audio User Interface, 21st British HCI Group Annual Conference (HCI 2007). Lancaster, UK.

Frauenberger, C., Stockman, T., & Bourguet, M.-L. (2007b). Pattern design in the context space: A methodological framework for auditory display design. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 513-518), Montreal, Canada.

Frysinger, S. P. (2005). A brief history of auditory data representation to the 1980s. *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction*, 4(1), 67-94.

Gaver, W. W., Smith, R. B., & O'Shea, T. (1991). Effective sounds in complex systems: The ARKola simulation. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 91)*, New Orleans, LA.

Gerber, E. (2003). The benefits of and barriers to computer use for individuals who are visually impaired. *Journal of Visual Impairment & Blindness*, September, 536-550.

Global music - The world by ear. (2006). [Concert]. London, UK: *International Conference on Auditory Display (ICAD2006)*.

Golledge, R. G., Loomis, J. M., Klatzky, R. L., Flury, A., & Yang, X. L. (1991). Designing a personal guidance system to aid navigation without sight: Progress on th eGIS component. *International Journal of Geographical Information Systems*, 5, 373-395.

Griffith, D. (1990). Computer access for persons who are blind or visually impaired: Human factors issues. *Human Factors*, 32(4), 467-475.

Grudin, J. (2001). Partitioning digital worlds: Focal and peripheral awareness in multiple monitor use. *Proceedings of the 2001 SIGCHI conference on Human Factors in Computing Systems (CHI 01)* (pp. 458-465), Seattle, WA.

Haas, E. C., & Edworthy, J. (1996). Designing urgency into auditory warnings using pitch, speed, and loudness. *Computing & Control Engineering Journal*, August, 193-198.

Hartmann, W. M. (1997). *Sounds, signals, and sensation: Modern acoustics and signal processing*. New York: Springer Verlag.

Hereford, J., & Winn, W. (1994). Non-speech sound in human-computer interaction: A review and design guidelines. *Journal of Educational Computer Research*, 11, 211-233.

Hermann, T., & Hunt, A. (2005). An introduction to interactive sonification. *IEEE Multimedia*, 12(2), 20-24.

Hetzler, S. M., & Tardiff, R. M. (2006). Two tools for intergrating sonification into calculus intstruction. *Proceedings of the International Conference on Auditory Display (ICAD2006)* (pp. 281-284), London, UK.

Jeon, J. Y., & Fricke, F. R. (1997). Duration of perceived and performed sounds. *Psychology of Music*, 25(1), 70-83.

Keller, P., & Stevens, C. (2004). Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied*, 10(1), 3-12.

Kramer, G. (1994). An introduction to auditory display. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces* (pp. 1-78). Reading, MA: Addison Wesley.

Kramer, G., Walker, B. N., Bonebright, T., Cook, P., Flowers, J., Miner, N., et al. (1999). The Sonification Report: Status of the Field and Research Agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display. Santa Fe, NM: International Community for Auditory Display (ICAD).

Lacherez, P., Seah, E. L., & Sanderson, P. M. (2007). Overlapping melodic alarms are almost indiscriminable. *Human Factors*, 49(4), 637-645.

Liljedahl, M., Papworth, N., & Lindberg, S. (2007). Beowulf: A game experience built on sound effects. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 102-106), Montreal, Canada.

Listening to the mind listening: Concert of sonifications at the Sydney Opera House. (2004). [Concert]. Sydney, Australia: *International Conference on Auditory Display (ICAD2004).*

Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (1993). Personal guidance system for the visually impaired using GPS, GIS, and VR technologies. *Proceedings of the First Annual International Conference, Virtual Reality and Persons with Disabilities*, San Francisco, CA.

Loomis, J. M., Marston, J. R., Golledge, R. G., & Klatzky, R. L. (2005). Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance. *Journal of Visual Impairment & Blindness*, 99, 219-232.

Mankoff, J., Fait, H., & Tran, T. (2005). Is your web page accessible? A comparative study of methods for assessing web page accessibility for the blind. *Proceedings of the ACM CHI Conference on Human Factors Computing Systems (CHI 05)*, Portland, OR.

Mauney, B. S., & Walker, B. N. (2004). Creating functional and livable soundscapes for peripheral monitoring of dynamic data. *Proceedings of the International Conference on Auditory Display (ICAD2004)*, Sydney, Australia.

Mauney, L. M. (2006). Individual differences in cognitive, musical, and perceptual abilities. . Unpublished master's thesis, Georgia Institute of Technology, Atlanta, GA.

McDonald, J. J., Teder-Salejarvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, 407, 906-908.

McGookin, D. K., & Brewster, S. (2004). Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Transactions on Applied Perception*, 1, 130-150.

McGuire, J. M., Scott, S. S., & Shaw, S. F. (2006). Universal design and its applications in educational environments. *Remedial and Special Education*, 3, 166-175.

Meyer, A., & Rose, D. H. (2000). Universal design for individual differences. *Educational Leadership*, 58(3), 39-43.

Mondor, T. A., & Amirault, K. (1998). Effect of same- and different-modality spatial cues on auditory and visual target identification. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 745-755.

Morley, S., Petrie, H., O'Neill, A.-M., & McNally, P. (1999). Auditory navigation in hyperspace: Design and evaluation of a non-visual hypermedia system for blind users. *Behaviour & Information Technology*, 18(1), 18-26.

Mynatt, E. D. (1997). Transforming graphical interfaces into auditory interfaces for blind users. *Human-Computer Interaction*, 12, 7-45.

Mynatt, E. D., & Edwards, W. K. (1992). Mapping GUIs to auditory interfaces. *Proceedings of the 5th Annual ACM Symposium on User Interface Software and Technology* (pp. 61-79), Monterey, CA.

Nees, M. A., & Walker, B. N. (2007). Listener, task, and auditory graph: Toward a conceptual model of auditory graph comprehension. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 266-273), Montreal, Canada.

Nees, M. A., & Walker, B. N. (in press). Data density and trend reversals in auditory graphs: Effects on point estimation and trend identification tasks. *ACM Transactions on Applied Perception*.

Neuhoff, J. G., Kramer, G., & Wayand, J. (2002). Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology: Applied*, 8(1), 17-25.

Palladino, D., & Walker, B. N. (2007). Learning rates for auditory menus enhanced with spearcons versus earcons. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 274-279), Montreal, Canada.

Peres, S. C., & Lane, D. M. (2003). Sonification of statistical graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)*, Boston, MA.

Peres, S. C., & Lane, D. M. (2005). Auditory graphs: The effects of redundant dimensions and divided attention. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 169-174), Limerick, Ireland.

Quinn, M. (2001). Research set to music: The climate symphony and other sonifications of ice core, radar, DNA, seismic, and solar wind data. *Proceedings of the International Conference on Auditory Display (ICAD2001)*, Espoo, Finland.

Ramloll, R., Brewster, S., Yu, W., & Riedel, B. (2001). Using non-speech audio sounds to improve access to 2D tabular numerical information for visually impaired uses. *Proceedings of the IHM-HCI 2001* (pp. 515-530), Lille, France.

Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G. P., et al. (2004). Global data on visual impairment in the year 2002. *Bulletin of the World Health Organization*, 82(11), 844-851.

Roth, P., Kamel, H., Petrucci, L., & Pun, T. (2002). A comparison of three nonvisual methods for presenting scientific graphs. *Journal of Visual Impairment & Blindness*, 96(6), 420-428.

Salmen, J. P. (1998). *Everyone's Welcome: The Americans with Disabilities Act and Museums*. Washington, DC: American Association of Museums.

Sandor, A., & Lane, D. M. (2003). Sonification of absolute values with single and multiple dimensions. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 243-246), Boston, MA.

Schmitt, M., Postma, A., & De Haan, E. (2000). Interactions between exogenous auditory and visual spatial attention. *Quarterly Journal of Experimental Psychology*, 53A(1), 105-130.

Smith, D. R., & Walker, B. N. (2005). Effects of auditory context cues and training on performance of a point estimation sonification task. *Applied Cognitive Psychology*, 19(8), 1065-1087.

Spence, C., & Driver, J. (1997). Audiovisual links in attention: Implications for interface design. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics Vol. 2: Job Design and Product Design* (pp. 185-192). Hampshire: Ashgate Publishing.

Spence, C., McDonald, J. J., & Driver, J. (2004). Exogenous spatial-cuing studies of human cross-modal attention and multisensory intergration. In C. Spence & J. Driver (Eds.), *Cross-modal space and cross-modal attention* (pp. 277-320). Oxford: Oxford University Press.

Stanley, R., & Walker, B. N. (2006). Lateralization of sounds using bone-conduction headsets. *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES 2006)* (pp. 1571-1575), San Francisco, CA.

Stevens, S. S. (1936). A scale for the measurement of a psychological magnitude: Loudness. *Psychological Review*, 43, 405-416.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A Scale for the Measurement of the Psychological Magnitude Pitch. *The Journal of the Acoustical Society of America*, 8, 185.

Stockman, T., Hind, G., & Frauenberger, C. (2005). Interactive sonification of spreadsheets. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 134-139), Limerick, Ireland.

Stockman, T., Rajgor, N., Metatla, O., & Harrar, L. (2007). The design of interactive audio soccer. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 526-529), Montreal, Canada.

Targett, S., & Fernstrom, M. (2003). Audio games: Fun for all? All for fun? *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 216-219), Boston, MA.

Tobias, J. (2003). Information technology and universal design: An agenda for accessible technology. *Journal of Visual Impairment & Blindness*, October, 592-601.

Turnbull, W. W. (1944). Pitch discrimination as a function of tonal duration. *Journal of Experimental Psychology*, 34, 302-316.

Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied*, 8, 211-221.

Walker, B. N. (2007). Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Applied Cognitive Psychology*, 21, 579-599.

Walker, B. N., & Cothran, J. T. (2003). Sonification Sandbox: A graphical toolkit for auditory graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 161-163), Boston, MA.

Walker, B. N., Godfrey, M. T., Orlosky, J. E., Bruce, C., & Sanford, J. (2006a). Aquarium sonification: Soundscapes for accessible dynamic informal learning environments. *Proceedings of the International Conference on Auditory Display (ICAD2006)* (pp. 238-241), London, UK.

Walker, B. N., Kim, J., & Pendse, A. (2007). Musical soundscapes for an accessible aquarium: Bringing dynamic exhibits to the visually impaired. *Proceedings of the International Computer Music Conference (ICMC 2007)* (pp. TBD), Copenhagen, Denmark.

Walker, B. N., & Kramer, G. (1996). Human factors and the acoustic ecology: Considerations for multimedia audio design. *Proceedings of the Audio Engineering Society 101st Convention*, Los Angeles.

Walker, B. N., & Kramer, G. (2004). Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. In J. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 150-175). New York: Academic Press.

Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays: An experimental assessment. *ACM Transactions on Applied Perception*, 2(4), 407-412.

Walker, B. N., & Kramer, G. (2006a). Auditory displays, alarms, and auditory interfaces. In W. Karwowski (Ed.), *International Encyclopedia of Ergonomics and Human Factors* (2nd ed., pp. 1021-1025). New York: CRC Press.

Walker, B. N., & Kramer, G. (2006b). Sonification. In W. Karwowski (Ed.), *International Encyclopedia of Ergonomics and Human Factors* (2nd ed., pp. 1254-1256). New York: CRC Press.

Walker, B. N., & Lane, D. M. (2001). Psychophysical scaling of sonification mappings: A comparision of visually impaired and sighted listeners. *Proceedings of the International Conference on Auditory Display (ICAD2001)* (pp. 90-94), Espoo, Finland.

Walker, B. N., & Lindsay, J. (2005a). The effect of a speech discrimination task on navigation in a virtual environment. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1538-1541), San Francisco, CA.

Walker, B. N., & Lindsay, J. (2005b). Using virtual environments to prototype auditory navigation displays. *Assistive Technology*, 17, 72-81.

Walker, B. N., & Lindsay, J. (2006). Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice. *Human Factors*, 48(2), 265-278.

Walker, B. N., Lindsay, J., & Godfrey, J. (2004). The Audio Abacus: Representing a wide range of values with accuracy and precision. *Proceedings of the International Conference on Auditory Display (ICAD2004)*, Sydney, Australia.

Walker, B. N., & Mauney, L. M. (2004). Individual differences, cognitive abilities, and the interpretation of auditory graphs. *Proceedings of the International Conference on Auditory Display (ICAD2004)*, Sydney, Australia.

Walker, B. N., Nance, A., & Lindsay, J. (2006b). Spearcons: Speech-based earcons improve navigation performance in auditory menus. *Proceedings of the International Conference on Auditory Display (ICAD2006)* (pp. 95-98), London, UK.

Walker, B. N., & Nees, M. A. (2005a). An agenda for research and development of multimodal graphs. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 428-432), Limerick, Ireland.

Walker, B. N., & Nees, M. A. (2005b). Brief training for performance of a point estimation sonification task. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 276-279), Limerick, Ireland.

Walker, B. N., & Nees, M. A. (in press). Theory of sonification. In T. Hermann, A. Hunt & J. Neuhoff (Eds.), *Principles of Sonification: An Introduction to Auditory Display*.

Walker, B. N., & Stanley, R. (2005). Thresholds of audibility for bone-conduction headsets. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 218-222), Limerick, Ireland.

Watson, C. S., & Kidd, G. R. (1994). Factors in the design of effective auditory displays. *Proceedings of the International Conference on Auditory Display (ICAD1994)*, Sante Fe, NM.

Wightman, F. L., & Kistler, D. J. (1983). Headphone simulation of free-field listening I: Stimulus synthesis. *Journal of the Acoustical Society of America*, 85, 858-867.

Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free-field listening. II. Psychophysical validation. *Journal of the Acoustical Society of America*, 85, 868-878.

Wilson, J., Walker, B. N., Lindsay, J., Cambias, C., & Dellaert, F. (2007). SWAN: System for wearable audio navigation. *Proceedings of the 11th International Syposium on Wearable Computers (ISWC 2007)*, Boston, MA.

Winberg, F., & Hellstrom, S. O. (2001). Qualitative aspects of auditory direct manipulation: A case study of the Towers of Hanoi. *Proceedings of the International Conference on Auditory Display (ICAD2001)* (pp. 16-20), Espoo, Finland.

Winberg, F., & Hellstrom, S. O. (2003). Designing accessible auditory drag and drop. *Proceedings of the the 2003 Conference on Universal Usability (CUU2003)* Vancouver, Canada.

Yalla, P., & Walker, B. N. (2007). Advanced multimodal menus for user interfaces: Georgia Institute of Technology Technical Report Number TBD.