Auditory Displays Article xx: Auditory Displays, Alarms, and Auditory Interfaces

Bruce N. Walker and Gregory Kramer*[†]
*School of Psychology and College of Computing, Georgia Institute of Technology, 654 Cherry Street, Atlanta, Georgia, 30332–1070 USA
[†]Metta Foundation, 310 NW Brynwood Lane, Portland, OR 97229, USA

1 INTRODUCTION

The world we hear is completely different from the world we experience through other sensory modalities. There are many features of the auditory world that make sound an excellent means of conveying information to the user of a system. Sound may be used in traditional alarms, in process monitoring displays, in computer system interfaces, and even to present complex multivariate data. It should be made clear that we are talking about the *intentional* (as opposed to *incidental*) use of sound. That is, sounds that are designed as part of the user interface of a system, as opposed to sounds that are unintentionally part of a system, but still convey information to a listener (e.g. the syncopated rhythm of an engine that indicates it is misfiring). Furthermore, for the most part we are considering only nonspeech audio.

1.1 WHY USE SOUND IN AN INTERFACE?

Many users simply cannot see a visual display. For visually-disabled persons, workers who must move around in the job environment, or users whose eyes are busy with other elements of a task, an auditory display may be the most appropriate means of information display. Acoustic signals tend to elicit an alerting response, and can be detected more quickly than signals presented via other modalities. Furthermore, the omnidirectional nature of auditory perception allows a listener to obtain information from a sound, regardless of location or orientation. Our lack of "earlids" highlights the fact that we are always listening, whereas we can choose to stop looking. For these reasons sound has served extremely well in a wide variety of alarms.

Humans are sensitive to changes in a sound, detecting, for example, a 0.2% change in frequency in some cases. Minute changes in temporal variables, such as rhythm or transient sounds, are readily detected by the auditory system. This makes sound especially suitable for monitoring time-varying displays. More recently, in data display or sonification, the excellent pattern recognition capacities of the auditory system have begun to be exploited to discover patterns in complex, multivariate data sets (see chapter on Sonification). In addition to recognizing patterns, we are capable of detecting, separating, and following multiple acoustic sources simultaneously, making sound ideal for the display of several concurrent data sources.

2 BRIEF CHRONOLOGY OF SOUND IN INTERFACES

Nonspeech sound has long been used to convey information to listeners. From drums to gongs, bells, and horns, humans have relied on simple auditory displays that have generally been used to convey different degrees of danger. More recently, sophisticated machines like airplanes, and complex systems such as power plants, have made extensive use of caution and warning tones. However, the actual information conveyed with these sounds has mostly been limited to announcing which one of a handful of emergency types is present (e.g. fire, dangerous altitude, toxic spill, a leak in the system).

Within the last few decades, more sounds have been used to inform users about the actual details of an event or object. More recently, auditory displays have been used to monitor factory processes, signaling a change in one of several production parameters. Computers have added greatly to the range of informative auditory interface elements, including the status of an ongoing process like file downloading, activity in multiple computer processors, and notification of the receipt of electronic mail.

Auditory displays have recently begun to take advantage of advances in sound synthesis hardware and software. Such displays are also benefiting from the development of the theory of dynamic auditory perception and how such theory may inform efforts to display actual data. Geiger counters and sonar are two well-known successes. Early efforts even included a prototype cockpit display to allow pilots to fly without the use of visual dials and gauges. Successful recent applications have presented spectrographic data to blind chemists, conveyed complex and dynamic patient data to anesthesiologists, displayed changing stock market prices, and have allowed seismologists to distinguish earthquakes from subterranean nuclear explosions (see Kramer *et al*. 1999). This type of auditory display is discussed further in the Sonification chapter.

3 TAXONOMY AND DEFINITIONS

It is helpful to consider auditory displays in terms of the symbolic–analogic continuum. Symbolic displays establish a mapping between a sound and an intended meaning, with no intrinsic relationship existing. Speech is explicitly symbolic, as are most auditory warnings and alarms. Analogic displays contain an immediate and intrinsic relationship between the display dimension and the information that is being conveyed. Using the tempo of a repeating tone to represent the heart-rate of a patient is highly analogic; that is, a change in heart rate is directly related, by analogy, to changes in tone presentation rate. Since this is a continuum, most auditory displays are somewhere between the extremes. Examples range from symbolic "earcons," which are families of event notifications (indicating, for example, that a computer file was opened); slightly more analogic auditory icons, which are event notifiers that also reflect the details of the events (indicating, for example, the type and size of computer file that was opened); and quite analogic representations of blood pressure or three-dimensional (3D) fluid dynamics. It is interesting to note that, historically, sound was used in a largely symbolic manner, but this has shifted more and more toward the analogic end of the spectrum (see Figure 1).

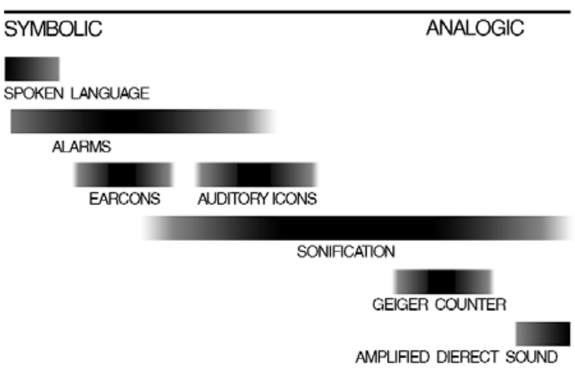

FIGURE 1 Relative positioning of various types of auditory displays along the symbolic-analogic continuum.

3.1 Definitions

Definitions are not yet solid in this nascent field, but they are becoming more so. Amongst symbolic displays (also the oldest end of the spectrum), the definitions are more stable. At the newer, more analogic end, the definitions are still evolving. The definitions we present here are becoming commonly used within the International Community for Auditory Display (www.ICAD.org).

Auditory display is a generic term including all intentional, nonspeech audio that is designed to transmit information between a system and a user.

Auditory alarm refers to a sound that alerts a user and conveys information about an event. Alarms are usually categorical and symbolic in nature.

Auditory interface indicates the use of sound for part or all of the display of system variables and other information, usually including information about the details of the objects involved or the underlying process. This term covers a wide range in the middle of the symbolic–analogic continuum, including multimedia computer interfaces.

Sonification is the use of nonspeech audio to present data. Specifically, sonification is the transformation of data relations into auditory relations, for the purpose of studying and interpreting the data. Sonification is at the more analogic — and modern — end of our conceptual continuum. See the Sonification chapter for more details on this topic.

4 WHEN TO USE SOUND IN AN INTERFACE

Sound and auditory displays can be used as a supplement to other (usually visual) displays, as an equal part of a multimodal display, or certainly as a complete display on its own. But the standard Human Factors practice of considering the users, their work environment, and the tasks that they need to accomplish is just as critical here as it is in the design of any display. Thus, sound may be a potentially very effective display device in a large number of situations where information must be conveyed, so long as due care is taken in the total information display design, and each modality is used appropriately. Many standard guides list the following circumstances where sound is the preferred display medium (e.g. Sanders and McCormick 1993): When the original signal is, itself, sound; the information to be displayed is short and simple; the information will not be referred to later; the information deals with events that change in time (key for process control and sonification); the information calls for immediate action (warnings, alarms); display real estate, illumination, or other reasons make visual display unacceptable; the receiver is moving around a lot; or when a verbal response is required.

Due to the evolution of the science of auditory display, sound is also recommended when multiple complex data sources need to be monitored or compared, and when the information contains 3D aspects that can be communicated using spatialized sound. Recommendations (e.g. Kramer 1994) also point to the use of sound for increases in perceived quality in a system, enhanced realism, wider options for enhanced learning and creativity, and heightened intermodal correlations and synesthesia. Many criticisms that have been made about auditory displays in the past are no longer valid, having been raised against relatively new and unsophisticated auditory displays, or sound hardware and software of limited capabilities. Well-designed auditory displays need not be annoying, disturbing, or extravagant, and can serve well in a broad range of display scenarios.

Auditory displays do have their downsides, of course. There is limited resolution for some variables, especially when represented with spectral or spatial sound attributes (though other attributes like pitch and loudness can show good resolution). While it need not be the case, many auditory displays lack an axis or reference points. It is generally difficult to discern absolute values with such displays. Thus, auditory equivalents of axes, tick marks, and labels should be included in any auditory display. Since we cannot "close our ears," sound may interfere with speech or other cognitive processing. In addition, since auditory displays are not bound by "line of sight," sound can be a problem in open work environments. Auditory displays in noisy environments, such as some factory floors, may be impractical not because of the user or the task, but simply because of masking problems in the environment. In such cases, use of the visual modality may be more appropriate. Another downside to auditory displays in some cases is the "absence of persistence," which means that sounds are not continuously present for study and comparison. In a practical light, it remains difficult to distribute sounds: it is hard to include sounds with printed material, and file size and format standards still pose a challenge for digital sound.

5 AUDITORY ALARMS

The most significant use of auditory alarms is on the flight deck of modern commercial and military aircraft. Ensuring audibility without disrupting subsequent speech communication is the basic goal in this situation. There is a vast array of possible alarm situations, so care must be taken in determining the actual conditions that will elicit an auditory alarm. To date, most cockpit alarms are simple alerts, with visual or speech supplements to specify the alarm condition. Patterson (1982) has provided detailed guidance in designing auditory alarms for cockpit scenarios (see below).

In many respects, control rooms can be considered very similar environments to cockpits, except that there will typically be more listeners in a control room, many of whom are moving around in a large space. Further, a larger range of warnings needs to be enunciated. Alarms, then, need to be separated into caution, warning, and danger categories, with appropriate sounds to communicate the category, specific type, and urgency of the problem (see Edworthy *et al.* 1991). Additionally, in a complex process control, cascading failures may produce a great number of fault conditions. This must not result in an incomprehensible cacophony of alarms, most of which will be immediately turned off if they seem to hinder the fault resolution task, rather than help it.

Many other auditory alarms are used. However, outside of the highly-trained populations of pilots and control room operators, ensuring recognition of different sounds, and the appropriate response action, poses a challenge to display designers. The goal of the alarm must be reduced to alerting the user of a problem status, then presenting problem details in other ways (speech, visual display). However, there is already a large number of fairly "standard" sounds that we are generally able to distinguish, ranging from telephones to police sirens to fire alarms. Within the constraints of population differences and learning requirements, alarm designers can maximize alarm perception and comprehension by taking user expectations and previous learning into consideration.

5.1 ALARM GOALS

The primary goal for an alarm is to alert, or notify the user of a certain event, time, or status, without disrupting the processing of information needed for the user's ongoing tasks. If an alarm is outside of the listener's hearing range, either in terms of pitch or loudness, or if it is masked by ambient noise, the alert will fail. At the same time, an alarm that is too loud may startle and confuse a user, and an alarm that occupies the same frequency range as other auditory information (most notably, speech) will disrupt perception of that source.

Beyond the simple alert, alarms should inform or convey information about the type of alarm situation. Alarms may convey higher levels of urgency based on sharper onsets, more rapid tempos, and louder presentation. (Edworthy *et al.* 1991) In addition, sophisticated alarms may present information about the details of the event, by varying the sound parameters. As an example, an alarm indicating an impending air traffic collision could spatialize the sound so that it appears to originate in the same place as the incoming airplane, thereby facilitating visual perception, and aiding appropriate control actions.

Once the alarm has alerted and informed the user, its job is not over. After notifying, the alarm should be designed to remind the user of the situation's continued alarm status. The alarm should help the user maintain situational awareness and not disrupt other activities. It is hoped that well-designed alarms will prove continuously helpful throughout an alarm situation, and will not simply be turned off because they aggravate or disrupt. However, the user must retain the ability to silence any or all auditory channels.

5.2 Alarm Design

In terms of loudness, Patterson (1982, 1990) recommends measuring the ambient noise levels during system operation, and plotting noise sound pressure level (SPL, measured in dB) against noise frequency (in Hz). The loudness of an alarm sound component at a given frequency should be set between 15 and 25 dB above the background noise at that frequency, but in no cases must it exceed 90 dB SPL. To avoid startle reactions, the onset and offset rates should not be too sudden, generally the range of 20–30 ms being acceptable.

For optimal perception, the alarm should include four or more frequency components, all within the range of about 1000–4000 Hz. The nature of the background noise spectrum, in particular the increased power and heightened masking effects of low-frequency noise, and the perception capabilities of the listening population need to be considered. Several frequency components will prevent an alarm from both masking, and being masked by, speech or other auditory information.

In general, an even spacing of frequency components will produce a more harmonic sounding alarm, one that tends to form a sonic "whole." The relative distribution of energy towards lower frequency components will tend to produce low-priority warnings, while shifting the energy to the higher frequency components will produce higher-priority warnings.

The tempo of a repeated pattern of sounds can affect detection, comprehensibility, disruption of other sounds, and perceived urgency of the alarm. Recommendations for the length of each pulse in a pattern are in the range of 100–150 ms, with pulses separated 50–300 ms apart, with tighter spacing leading to higher perceived urgency. As an example, a pattern of four 100 ms pulses separated by 100 ms intervals could be presented loudly at first, followed by digitized speech to specify the alarm details, and then repeated every few seconds at a lower loudness level, to avoid disruption of subsequent communication. Whole families of alarms can be constructed in this way, using variations of pitch, timbre and tempo patterns, not to mention the addition of digitized speech.

6 AUDITORY INTERFACES

Sound is used increasingly in human–machine interfaces to communicate more information than simple (symbolic) alarms, but still not as much as (analogic) data sonification. There are several common ways that sounds are being used in the interface. A common use is for status indicators and notifiers. These simple sounds often provide feedback for user input and actions, or provide notification of some event. For example, a "bop" sort of sound commonly indicates a prohibited action or an unacceptable entry. In an otherwise graphical user interface (GUI), sonic additions to widgets such as scroll bars and menus can assist users in making selections and choices more effectively than visual-only interface elements. Also, many sounds are used to indicate the completion of an action or the arrival of messages and electronic mail. Note, though, that all of these sounds are closely related to alarms, containing only minimal information about the specific event.

As discussed under Alarms, families of simple messages can be constructed by varying the spectral or temporal aspects of auditory notifiers. In the computer interface, the term "earcon" (a play on the word icon) describes small "languages" that can be made up of patterns of two or three musical notes. Consider an example where a three note motif composed of a long tone followed by two short tones indicates an action has been done to a computer file. A rising pattern, like "bong-bing-bing" might mean the file was opened, whereas a descending pattern like "bing-bong-bong" could mean the file was closed. Different timbres could represent different file types. Clearly this is a symbolic, abstract use of sound, but plenty of information can be layered into the sound families. A major limitation is the training that is required for naive listeners to learn the symbology.

As an alternative to earcons, auditory icons use sounds that in some way represent the actual features of objects involved in actions. For example, if a computer's storage device is represented by a sound, that sound can vary depending on the type of device, the capacity of the device, and the available storage space. Alternatively, a sound that is used to display the status of a file transfer procedure can sound like a bottle filling with water, with the perceived "size" of the container and "composition" of the liquid being determined by the size and type of file that is being transferred. In this way, the auditory icon is parameterized, and shares a more direct, analogic relationship with the object and the event it is representing.

Recent developments in computing now allow auditory display designers to make a sound appear to originate in a specific position in three-dimensional (3D) auditory space. This is known as spatialization of sound. While the success of the technique depends on a number of factors, and is not perfect in any case, 3D sound can be used to communicate information to a listener in a number of important ways: first, if the information that is being communicated contains inherent spatial information (such as the position of a tumor in a section of tissue), then 3D sound can very directly represent this to the user. On the other hand, the location of a sound in the virtual 3D world can be used to symbolically represent any sort of information. For example, the pressure of a gas might be represented by the left-to-right location of a sound, while the temperature of the gas might be represented by the vertical height of the sound source. These might seem like intuitive or natural mappings, but they are still largely symbolic uses for spatialized sound. For more on this topic, see Spatial Auditory Displays.

7 RECOMMENDATIONS

Sound can be a powerful and effective information display option, creating engaging and compelling displays. Auditory alarms have proven highly useful in a wide range of situations. However, as with all interface design, failure to consider the user, the context, and the requirements of the task can lead to annoying or intrusive auditory displays. For successful auditory interfaces, a solid understanding of both technical and perceptual aspects of the auditory medium is required. Expertise in sound design and synthesis is recommended for any auditory display, and especially for more sophisticated sonification applications. All auditory displays need to be thoroughly tested for comprehension and usability with representative listeners, especially in situations where sound has not previously been employed. What may sound "intuitive" to a designer may not be so for end users. The scientific investigation, design, and validation of using sound to display information is a relatively new field. Developments to date point to a promising future for auditory displays.

REFERENCES

EDWORTHY, J., LOXLEY, S. and DENNIS, I., 1991, Improving auditory warning design: relationship between warning sound parameters and perceived urgency. *Human Factors*, 33(2), 205–31.

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

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)). AQ1

PATTERSON, R.D., 1982, *Guidelines for Auditory Warning Systems on Civil Aircraft* (London: Civil Aviation Authority).

PATTERSON, R.D., 1990, Auditory warning sounds in the work environment. *Philosophical Transactions of the Royal Society of London*, B327, 485–94.

SANDERS, M.S. and MCCORMICK, E.J., 1993, *Human Factors in Engineering and Design* (7th ed.) (New York: McGraw-Hill).