# Dissonance on audio interfaces

By Marti A. Hearst
Xerox Palo Alto Research Center
hearst@parc.xerox.com

An important aspect of a complex intelligent system is the human-computer interface. Although most of our readers are well-acquainted with graphical user interfaces, in this installment we discuss the perhaps less-familiar topic of audio interfaces. The controversy surrounds which of three competing audio interface approaches is most effective: sonification (Stephen Barass), earcons (Stephen Brewster), and auditory icons (Beth Mynatt).

To help facilitate comparison, each audio expert has described a design using his or her approach for an example problem, created specifically for this discussion by Michael Albers (see http://computer.org/expert/ex1997/extras/ for the audio files our contributors created). Albers also wraps up the discussion with a brief comparison of the results.

I would like to acknowledge Beth Mynatt for organizing the contributions for this issue.

***—Marti Hearst***

## Setting the stage

*Michael C. Albers, Sun Microsystems*

The human operators of complex systems, such as computer-integrated manufacturing systems, power plants, and aircraft flight decks, are highly educated and trained. They can monitor and manage their particular systems under normal conditions and during system malfunctions. These operators typically must deal with an enormous quantity of data conveying system health that is presented to them on multiple computer screens.[1,2]

To assist with this data avalanche, automation's role in the control of these complex systems has expanded. The increase of automation changes the operator's role from a controller to a monitor and, when needed, a troubleshooter responsible for fault detection, diagnosis, and compensation.[3] Although automation has shifted operator responsibilities, there is no substitute for human decision making and experience to set high-level system goals, monitor system states, and compensate for anomalies that the automation systems were not designed to handle.[4]

When anomalies occur for which automation cannot compensate, these operators must quickly diagnose and correct the anomalies. They identify faulty system components by observing incorrect systems states, creating hypotheses based on their experience, and testing each hypothesis by manipulating the system. The interactions between, dynamics of, and sheer number of elements in these systems complicate this method of anomaly diagnosis and compensation. The huge number of system components increases the alternatives available to explain system anomalies. Furthermore, determining which anomalies are causes and which are effects is difficult because any single anomaly quickly proliferates through the systems to cause other anomalies.

Steam-propulsion plants are prototypical examples of complex engineering systems. Engineers have studied them extensively and developed simulations of them to assist training about and the control of these complex engineering systems. Steam-propulsion systems deconstruct into subsystems such as fuel-oil, lubrication-oil, main-steam, auxiliary-steam, cooling, and turbine. Likewise, each subsystem (the main-steam system, for example) decomposes into subsubsystems, such as air supply, economizer, and boiler. Finally, these subsubsystems are composed of physical components, which form physical interconnections to create a network of interactions and failure points.

This decomposition of the system parallels the operators' knowledge. Generally, plant operators and system designers think of higher-level pieces of the system in hierarchical, functional terms, whereas they understand lower levels in terms of physical interconnections. For example, they think of the main-steam subsystem as a heat-transfer unit, but they consider the boiler as an interconnected collection of physical components.

### Cascading anomalies

As an example of how single anomalies cascade through these systems, consider a hypothetical (and simplistic) marine steam-power plant. The emergency cooling-water valve has been sporadically faulty since the power plant went online. Consequently, the operators have never trusted this component's readings, have shut off its visual and auditory alarms, and have blocked its data from being used by the plant's automation system. (See Figure 1 and also the "Definitions" sidebar.)

At time index zero (T+0:00), the ECWV fails mechanically and slowly starts to close (see Table 1). Accordingly, the amount of cooling water flowing to the boiler slowly decreases. Soon (T+1:15), the boiler's hull temperature starts to rise. To compensate for the rising boiler-hull temperature, the plant's automation system starts opening the main cooling-water valve until it is completely open (T+2:10). Because the ECWV is blocking the cooling water's flow, opening the main cooling-water valve has little effect.

Because the boiler's hull temperature is still inching upward (at T+2:35), the amount of fuel water in the boiler is decreasing, while the amount of steam to power the turbine, the power produced by the turbine, and the amount of lubrication oil flowing to the turbine all are increasing. At T+2:55, in an attempt to stabilize this developing situation, the automation system takes the following steps. To compensate for the declining fuel-water level in the boiler, the automation system slowly opens the valve controlling the amount of fuel water fed to the boiler. It slowly opens the steam vents in the turbine to vent steam. Finally, to compensate for the boiler hull's increasing temperature, it slowly decreases the amount of fuel oil fed to the burners.

Table 1. Consequences of a failing emergency cooling-water valve in a hypothetical marine steam-power plant.

| | T+0:00 | T+1:15 | T+2:10 | T+2:35 | T+2:55 | T+5:15 | T+5:20 | T+5:25 | T+5:30 | T+8:30 |
|---|---|---|---|---|---|---|---|---|---|---|
| ECWV | Failing | | | | | | Failed | | | |
| MCWV | | | Opening | | | | | | | |
| FWFV | | | | | Opening | | | | | |
| FOFV | | | | | Closing | | | | Closing | Closed |
| TSV | | | | | Opening | | | | | Open |
| CWL | Decreasing | | | | | Dangerous | | | | |
| FWL | | | | Decreasing | | | | Dangerous | | |
| LOL | | | | Increasing | | | | | | |
| BHT | | Increasing | | | | | | Dangerous | | |
| TSL | | | | Increasing | | | | | | |
| TRPM | | | | Increasing | | | | Dangerous | | |
| SCWR | | | | | | | | Water detected | | |

These steps by the automation system help stabilize the system in the short run. However, within a few minutes, the automation system's corrective steps cannot overcome the fact that the amount of cooling water flowing to the boiler is inadequate. At T+5:15, the amount of cooling water in the boiler reaches a *dangerous* level. (All system components have levels at which they are considered to be in a dangerous state.) The ECWV has completely failed at T+5:20, restricting any more cooling water from reaching the boiler. As the remaining cooling water in the boiler heats up and boils off, the single component failure starts to quickly cascade through the system. At T+5:25, the boiler's hull temperature, the fuel-water levels, and the turbine's RPMs reach a dangerous level. At T+5:30, in a last-ditch effort to keep the boiler from melting down, the plant's automation system dramatically scales back the amount of fuel oil fed to the burners. By T+8:30, the automation system must completely shut off the fuel oil to the burners and vent any remaining steam from the system.

If the operators of this marine steam-power plant cannot diagnose and correct the true problem by T+8:30, the ship will be relying on backup power for everything that needs electricity (radios, medical, and navigation, for example). Failures or shutdowns of complex systems, including this power plant, are extremely costly in monetary and human terms. If the operators could correctly identify the ECWV as the faulty component, they could avert the power plant's shutdown by going to the ECWV's physical location and manually opening it.

## Adding sound to human-machine interfaces

Complex system monitoring and anomaly diagnosis by operators occur through a workstation with a graphical user interface. Most workstation environments place a heavy workload on the visual systems of operators in complex work environments.[1,2] This heavy visual load often makes these complex systems extremely difficult to manage.[2]

Using different human senses in the computer interface lessens the visual load on an operator. One way to lessen the visual load is to add sound "so that operators may hear rather than see displays."[1] Adding sound to human-machine interfaces is especially useful in complex dynamic systems in which users cannot be expected to notice all visual feedback.[5] Furthermore, time-varying data, multivariate data, background processes, and transient conditions—common elements of complex engineering systems—are well suited to sonic representation.[6] Operators benefit from the addition of an auditory interface to a graphical interface for a complex system.

The following three essays outline approaches to the inclusion of audio to aid the operators of the complex system in the failure scenario I've just outlined.

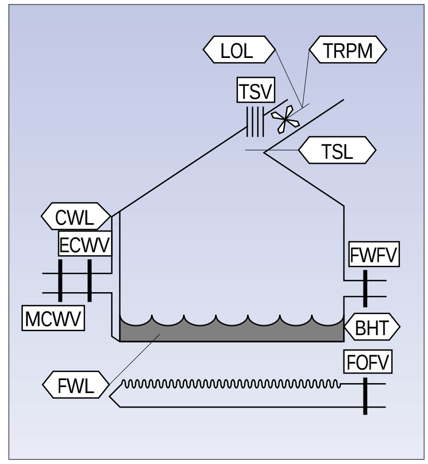



Figure 1. Hypothetical marine steam-power plant with faulty emergency cooling-water valve: ovals represent temperatures and levels; boxes represent valves.

### Definitions

| | |
|---|---|
| BHT | Boiler-hull temperature |
| CWL | Cooling-water level |
| ECWV | Emergency cooling-water valve |
| FOFV | Fuel-oil feed valve |
| FWFV | Fuel-water feed valve |
| FWL | Fuel-water level |
| LOL | Lubrication-oil level |
| MCWV | Main cooling-water valve |
| SCWR | Spillover cooling-water reservoir |
| TRPM | Turbine rpms |
| TSL | Turbine steam level |
| TSV | Turbine steam vents |

## *Sonification*

*Stephen Barrass, CSIRO Mathematical and Information Sciences*

The challenge is to use only pure tones—the sound of a sine wave. This limited palette of sounds is a major design issue, but the task, data, and perception of the sounds are equally important. [6–8]

### Task

The task for the human operators is to monitor the system and diagnose anomalies that arise. Monitoring is a continuous task that can become tedious. Diagnosis is seldom necessary but is demanding and requires focused attention and concentrated problem solving.

Monitoring and diagnosis require different information. Monitoring is an ongoing question—"Is everything okay?"—that requires an answer about the whole system. Diagnosis is a discrete, local question—"Is this component okay?"—that requires a precise answer about a part of the system. Most displays used for monitoring are designed for diagnosis and show irrelevant data or *noise* that might adversely affect the monitoring task. Sometimes, this problem is fixed by an auditory alarm that alerts the operators to a dangerous condition that might go unnoticed in the visual display.

The alarm relieves the operator from constant, focused attention to a visual panel. However, the presence of an alarm changes the monitoring question from "Is everything okay?" to "Is there danger?" and might reduce awareness of the overall state. An alarm is always an emergency, so time for assessing the situation is short, and cool reasoning might be difficult. In the steam-propulsion scenario, the first alarm would sound only three minutes before shutdown, when the cooling-water level becomes dangerously low.

Instead of reacting to emergencies, it would be better to anticipate and prevent them. Perhaps we can design an auditory display that allows the operator to maintain awareness of the system and to begin diagnosis before an emergency develops.

## Data

To understand the data, we must understand the organization of the system and its subsystems. There are three main subsystems—cooler, boiler, and turbine—connected by physical variables CWL (coolant-water level), TSL (turbine steam level), and TRPM (turbine revs), as shown in Figure 2. The internal state of the coolant subsystem is measured by SCWR (coolant spillover), and the internal state of the boiler is measured by BHT (hull temperature) and FWL (fuel-water level). There are also control-feedback variables MCMW, ECWV, FOFV, FWFV, and TSV, which confirm that valves and vents in each subsystem are operating. The variables are all continuous-valued.

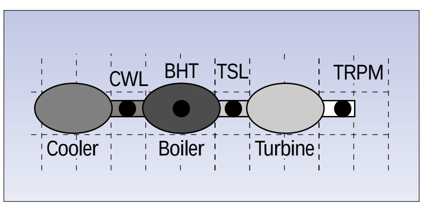



Figure 2. Useful information.

## Information requirements

A useful monitoring display will allow the operator to quickly, confidently, and correctly answer the question "Is everything okay?"

- yes, okay
- yes, but unusual
- yes, but something is going wrong
- no, danger

The answers depend on information about individual variables that have danger levels, and about the overall state of the system. The variables that have danger levels are CWL, BHT, TSL, and TRPM. If we combine these with the subsystem output variables that indicate the system state (CWL, TSL, and TRPM), we find four variables provide useful information for monitoring: CWL, BHT, TSL, and TRPM, as Figure 2 shows.

## Perception

The *streaming* theory explains listening by *primitive* and *schema* processes that group acoustic variations into distinct sounds.[9] The primitive process is a fast, preattentive, default grouping by factors such as spectral similarity and simultaneous onset. The slower schema process involves recognizing familiar patterns and explains the influence of attention and learning on what is heard.

A display that engages the primitive process suits a monitoring task because it is less affected by attention and training than a schema-based display. Factors that influence primitive grouping have been measured with the Van-Noorden *gallop* effect, generated by a pair of sounds, X and O, repeated in sequence, XOX-XOX.

If the sounds are perceptually similar, they group into the same stream and a distinctive galloping rhythm is heard, XXX-XXX. If X and O do not group, two distinctly separate rhythmic streams, X-X-X-X- and --O--O--, are heard, and you can choose to listen to one or the other. A list of factors that influence primitive grouping can guide the design of a display to engage the primitive process:

1. Onset intervals in the range 60–150 milliseconds
2. Difference between spectral centroids
3. Frequency difference from 4–13 semitones
4. Binaural harmonic correlation
5. Correlated frequency modulations
6. Correlated amplitude modulations
7. Harmonic relations
8. Parallel spectral movement
9. Synchronous onsets

## Display device

The constraint to a palette of pure tones can be viewed as a characteristic of the audio device that will be used to realize the display. An audio device like this is not uncommon in products that are limited by size, power, portability, maintenance, and cost. You might hear pure tones when your digital watch alarm goes off or when a mobile pager rings. Devices such as these play one tone at a time through a single speaker.

The parameters for producing sounds on such a device are amplitude, frequency, phase, and duration. Although it seems limited, this device can engage six of the nine primitive factors. Duration can engage factor 1. Pure tones do not have a centroid, so they cannot engage factor 2. Frequency can engage factor 3. A monaural device cannot engage factor 4. Correlated changes in amplitude and frequency can engage factors 5, 6, and 8. Harmonic relations in frequency can engage factor 7. A single-channel device cannot engage factor 9.

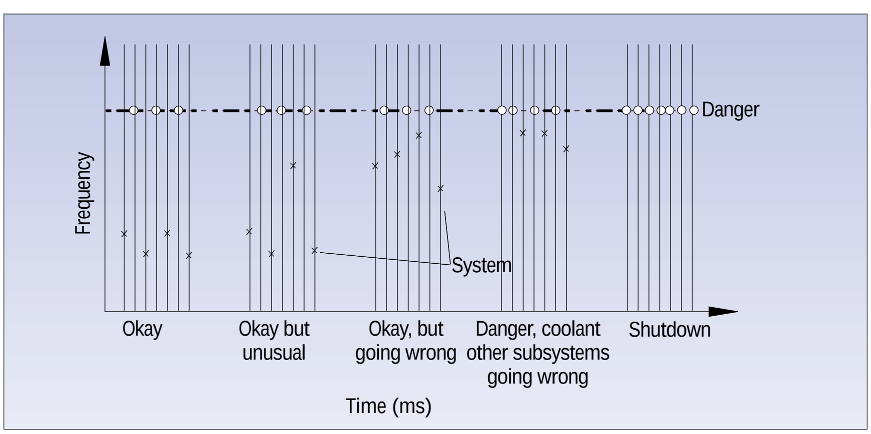



Figure 3. Galloping tone display.

## Information representation

The representation should allow the operator to hear useful information about CWL, BHT, TSL, and TRPM by primitive grouping of pure tones on the audio device. Each variable controls the frequency variation of a pure tone X. "Danger" for each variable maps to the same frequency O. The variable tones X are interleaved with "okay" tones at frequency O to form a gallop sequence XOXOXOX (see Figure 3). All tones are 100-ms long to engage streaming by factor 1. Tones closer than approximately six semitones in frequency group by factor 3. Under normal conditions, the "okay" tones are heard as a distinct O-O-O triplet that signifies "yes, okay." If any variable X approaches danger frequency O, the "okay" triplet becomes erratic, signifying that something is going wrong. Figure 3 shows the "okay" triplet by the three regularly spaced circles at the danger frequency. The figure illustrates each answer (okay, unusual, going wrong, danger, and shutdown). The X variables form a distinct second stream, which has a characteristic pattern shown by the circles below the danger line.

If a variable reaches danger, the "okay" triplet transforms to a "danger" gallop that signals which subsystem is involved—for example, CWL "danger" has rhythm OO-O-O. As more variables reach danger, the "danger" gallop becomes faster and more insistent. Meanwhile, the stream of variable tones that are not grouped with the gallop form a "system" stream that provides information about the rest of the system. The operator can learn the meaning of the system patterns (or schemas) and use them to diagnose the problem. For example, TRPM might usually track TSL, so these tones usually stream together. In an unusual condition, TSL might go high while TRPM remains low, causing a recognizable "unusual" pattern.

In the steam-propulsion scenario, you can hear that something unusual is happening soon after T=0:00, because the decrease in CWL can be heard as an unusual variation in the system pattern. This could prompt diagnosis right away, with eight minutes to spare. An experienced operator might diagnose that CWL is changing from this pattern. At T+1:15, the increase in BHT causes further unusual perturbations in the system stream. At T+2:35, the propagation of the anomaly into other subsystems causes a radical shift in the "system" pattern, and diagnosis should be well underway. In the fourth minute, CWL approaches dangerous levels and the "okay" gallop sounds unstable. At T+5:15, the "okay" gallop has transformed into a "danger" gallop with the signature OOO-O-O rhythm of danger in the coolant subsystem. The "danger" gallop builds in rate and insistence as each subsystem becomes dangerous, until at T+5:25, the only sound is a single urgent OOOOOOO stream that warns that the entire system is in danger.

## Summary

This auditory display uses only pure tones. The display supports the monitoring task that occupies most of a system operator's time. It improves on alarms by allowing the operator to maintain awareness of the system as a background activity, to detect unusual conditions, and to begin diagnosis before an emergency develops.

# Earcons

*Stephen Brewster, University of Glasgow*

Earcons are a method for presenting information about the state of the marine steam-power plant. Earcons are abstract, musical tones that can represent parts of an interface.[10] My detailed investigations show that earcons are an effective means of communicating information in sound.[11] I have also developed guidelines for their use that I used to design the sounds described here.[12]

Earcons are constructed from motives. These are short, rhythmic sequences that can be combined in various ways.[10] Earcons use many principles from music and psychoacoustics. They have successfully improved the usability of standard graphical-interface components such as buttons and scrollbars.[13] They have also provided navigation cues in telephone-based interfaces and interfaces for blind users.[14]

## Earcon design

As the earcon design guidelines suggest, I started with timbre (instrument), register, and rhythm.[12] Earcons should make up a musical whole. The operator will recognize the sound for the whole and so know when the system was running well. As things go wrong, parts of the sound will change, indicating where the problem is.

Simple rules defined the earcons for each system in the power plant to make the sounds easy to remember. These were

- timbre and stereo position define the system (cooling-water system or turbine system),
- rhythm and tempo define the level (increasing or decreasing), and
- pitch defines the state of a valve (opening or closing).

Because several earcons can play simultaneously, these attributes let the operator hear each system as a separate auditory source so that they will not be confused with each other.

I gave each of the six systems (cooling water, fuel water, fuel oil, turbine, boiler, and lubrication oil) an instrument (for example, the cooling-water system might use an organ and the turbine a violin) and a stereo position. The stereo position would match the position in the plant—for example, turbine on the right and cooling-water system on the left.

Rhythm and tempo indicate the level in the system (increasing/dangerous, decreasing/dangerous). A faster tempo indicates an increase in the level, and vice versa. There is a natural link between the concept of increasing and faster tempo that would make this mapping easy to remember. Reverb added to the earcons for decreasing/dangerous further differentiates them from the increasing/dangerous ones.

To indicate a dangerous level (either increasing or decreasing), I added a danger earcon to the mix. This bell timbre plays with the same rhythm and tempo and in the same stereo position as the system in the dangerous state, so the operator can hear which system is going wrong.

The pitch structure of the notes played indicates the state of the valve in a system (open, opening, closing, closed). Higher-pitched notes mean opening and lower-pitched ones closing. A three-note chord indicates that the valve is fully open (or closed). The notes played increase or decrease through the notes in the C major scale, so that earcons playing simultaneously sound harmonious. An enhancer effect applied to the

### Hear the music!

You'll find the audio files described in these essays at

http://computer.org/expert/ex1997/extras/

opening/open earcons also further differentiates them from the closing ones.

Therefore, by listening to an earcon's timbre/stereo, operators will know the system; by listening to its rhythm/tempo, they will understand the level in the system; and by listening to its pitch, they will know the state of the valve in the system.

To reduce the overall sound level (and avoid problems of annoyance), the earcon for any system that has not changed for a period of time will get quieter, which will help the plant operator to become habituated to it.


### Earcons applied to the marine power plant

Here are the results of applying earcons to the hypothetical test case.

*T+0:00.* The amount of cooling water flowing into the boiler (CWL) is decreasing. The tempo of the cooling-water earcon (played with an organ timbre in the left stereo position) starts to decrease to indicate the reduction in flow. The pitch indicates the state of the MCWV.

*T+1:15.* The BHT starts to rise. The boiler earcon—played with a violin timbre and a right stereo position—increases in tempo.

*T+2:10.* The MCWV opens. The cooling-water earcon rises in pitch. When the valve is fully open, a high-pitched chord sounds.

*T+2:35.* The BHT is still rising, signaled by a further increase in tempo. The FWL is dropping, so the fuel-water earcon's tempo slows. The TRPM level is increasing, so the turbine earcon tempo increases. The TSL is also rising. As the turbine has two level indicators (steam level and RPM level), we need another method for presenting the TSL. For this, a second earcon plays, directly after the main-turbine earcon, that uses the same instrument and stereo position (therefore, it is clearly associated with the turbine). It has a different rhythm to indicate that it is a different level indicator. This earcon has its own tempo that indicates the TSL. Finally, the LOL is increasing, so the lubrication oil earcon also plays at a faster tempo.

*T+2:55.* The FWFV slowly opens, indicated by a rise in pitch of the fuel-water earcon. The TSV opens, so the pitch of the turbine earcons increases. The FOFV starts closing, so the fuel-oil earcon decreases in pitch.

*T+5:15.* The CWL reaches a dangerous level. The alarm earcon mixes into the cooling-water earcon (which now plays at a very slow tempo).

*T+5:20.* The ECWV has failed, so no cooling water is flowing into the boiler. The cooling-water earcon plays at the slowest tempo, with the alarm earcon accompanying it.

*T+5:25.* The BHT plays at a fast tempo with the alarm earcon mixed into it. The FWL is dangerously low, so the earcon plays at a slow tempo, the alarm earcon mixed with it. The TRPM has reached a dangerously high level, so its earcon plays at a fast tempo with the alarm earcon mixed into it.

*T+5:30.* The FOFV continues closing, so the pitch of the fuel-oil earcon lowers.

*T+8:30.* The FOFV is closed completely, so the earcon plays with a low-pitched chord. Steam is vented (TSV), so the turbine earcons play with a high-pitched chord as the vent is fully open.

## *Auditory icons*

*Elizabeth D. Mynatt, Xerox PARC*

In this essay, I will describe using *auditory icons* as a method for presenting information about the marine steam-power plant. William W. Gaver introduced the concept of auditory icons based on the premise that people describe sounds relative to the objects interacting to make the sounds. Computers and other man-made devices make a variety of beeps, buzzes, and other artificial noises that otherwise we would never hear. Conversely, we typically hear things crumbling, sloshing, or colliding—we hear the results of objects interacting together. When asked to describe a sound, we tend to describe it in terms of the objects that generated the sound, such as a door slamming, stairs creaking, or glass breaking. In other words, we describe sounds in terms of their sources, not in classical terms such as pitch and duration, or in musical terms such as timbre or melody.[5]

What this realization means to interface designers is that we can use sounds to make users think of familiar objects in the same way that icons in visual interfaces characterize commonplace objects. Both the film and game industries provide engaging examples of sound-effects design where not only objects but qualities of those objects—the weight of a slamming door, the menace of an approaching animal, the heat of a fire—are made salient through sound. The best real-world example relevant to this design is that of a mechanic listening to a running car. The sounds of belts, fans, steam, and combustion can point to individual problems as well as form a cohesive "picture" of the system's health.

### Auditory icon design

For this marine power plant test case, I have given each of the subsystems a class of auditory icons for use in identifying the subsystem as a whole. In this design, the subsystem and sound groupings are

- Cooling-water levels as babbling brooks to raging streams
- Fuel-water levels as bubbling to boiling liquids
- Fuel-oil levels as crackling fire to raging inferno
- Boiler-hull temperature as meat sizzling
- Turbine-steam vents as slowly turning fans to rapid propellers

- Turbine-steam levels as incremental bursts of steam
- Turbine RPMs as small to large engines

These sounds connect to observable states of the system rather than connecting with the various controls. For example, when we drive a car, we rarely hear the state of controls—what gear we are in—but we hear the results of our actions and other influences. This design allows us to gear visual displays to depicting control states. Operators listen to the system, likely commenting to each other about the current state, and then while looking at the various monitors next to the controls, they individually adjust the system. This multimodal separation of information matches the existing practices of operators as they jointly diagnose a system and then partition control tasks.

Auditory icons are generally composed of sampled sounds, rather than the synthesized sounds used in the two previous designs. Although there are tools for generating pure tones and MIDI-controlled devices for generating musical scores, there are no tools for generating complex auditory icons such as the sound of a raging river or an outboard boat motor. Designers typically sample sounds from real-world sources and then add techniques for looping samples and transitioning between different samples. Surprisingly, designers rarely use the actual sound of the thing they want to represent. For example, Myst's designers commented on using the sound of driving over gravel to represent fire. In this essay, I'll refer to what the sounds should make you think of (water, fire, fans) without discussing their actual sources.

Without synthesizing auditory icons, representing a spectrum of states is difficult. For each subsystem, three samples will represent three states (normal, below threshold, and above threshold). To indicate approaching transitions, the timing of the looping sample will change. When all is normal, the pace of the looping samples will be synchronized, evoking the feel of a well-running system akin to a smoothly running car. As subsystems move to thresholds, the pace of the samples will fall out of sync, either lagging or jumping ahead of a silent but tangible clock for the system as a whole. After passing a threshold value, the subsystem's sample will change appropriately. The pacing of the sounds will continue. For example, a system running at half power would center on the below-threshold samples at a slower pace. Because timing serves as a discriminator of samples, in this design you would want the operator to "hear the loop," as opposed to typical background sounds in movies and games, where hearing the loop would be distracting and indicate a poor design.

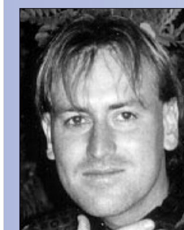
When a subsystem moves past a dangerous threshold, a brief alarm will precede the beginning of the looping sample. The alarms are simple alerts that will likely prompt the operators to use the visual displays to confirm dangerous states.

The auditory display serves primarily to give the operators a continuous, background feel for the system as a whole, as well as a feel for each subsystem. This background awareness is key to the quick handling of problems as they arise. The danger with automated systems is that while under automation the subsystems will move in and out of various threshold states without the awareness of the operators. Only when the system reaches a dangerous state and automation fails will the operator be engaged with comprehending the system's state. Awareness of the path from normal to dangerous is a clear benefit during diagnosis and recovery.

### Auditory icons applied to the marine power plant

Assuming that the system has reached a normal, steady state, the operator will hear six looping samples in sync: a small running river for cooling water, a medium boiling liquid with popping sounds (bubbles bursting) for fuel water, a medium fire (lots of crackling) for fuel oil, fans at medium speed for steam vents, a medium level of steam bursts for steam levels, and a medium-size outboard motor for the turbine.

**Michael C. Albers** is a user interface designer in Sun Microsystems' JavaSoft division. His research interests include human-computer interaction, particularly the use of sound, cognitive science, and the history of technology. He has a BS in cognitive science from the University of California, San Diego. He spent three years at the Georgia Institute of Technology developing auditory displays for complex engineering systems. He is a member of the ACM and ACM's Sigchi. Contact him at Sun Microsystems–JavaSoft, 901 San Antonio Rd., Palo Alto, CA 94303; michael.albers@sun.com; http://java.sun.com/people/mca/.



**Stephen Barrass** is a research scientist at the Digital Media Information Systems group at the CSIRO Mathematical and Information Sciences. His research interests include auditory display, human-computer interaction, multimedia content analysis, auditory scene analysis, algorithms for learning and mimicking sounds, algorithmic dance music, sound culture and art, and the listening behavior of animals. He has a BE from the University of South Wales and a PhD in information technology from the Australian National University. He is a member of the IEEE and the ACM. Contact him at CSIRO Mathematical and Information Sciences, GPO Box 664, Canberra Act 2601, Australia; stephen.barrass.cmis.csiro.au; ftp://ftp.cbr.dit.csiro.au/staff/stephen/stephen.html.

**Stephen Brewster** is a lecturer in human-computer interaction in the Department of Computing Science, University of Glasgow. His research interests include multimodal human-computer interaction and haptic interaction, particularly in all aspects of the use of sound to improve interaction in graphical computer interfaces and telephone-based systems. He has a BSc in computer science from the University of Hertfordshire and a PhD in human-computer interaction from the University of York. He is a member of the ACM, Sigchi, and the British Computer Society. Contact him at the Dept. of Computing Science, Univ. of Glasgow, Glasgow, G12 8QQ, UK; stephen@dcs.gla.ac.uk; http://www.dcs.gla.ac.uk/~stephen/.

**Elizabeth D. Mynatt** is a member of the research staff at the Xerox Palo Alto Research Center. Her research interests include supporting autonomous collaboration, designing network communities, and sonifying collaborative spaces bridging cyberspace and the physical world. She received her MS and PhD in computer science from the Georgia Institute of Technology. At Georgia Tech, she developed a method for transforming graphical user interfaces into auditory user interfaces. One application of this work is providing access to graphical interfaces for blind users. Contact her at Xerox PARC, 3333 Coyote Hill Rd., Palo Alto, CA 94304; mynatt@parc.xerox.com; http://www.parc.xerox.com/mynatt.

## In upcoming issues

- **The feasibility of online-only journals**
- **Bayesian networks**

*T+0:00.* As the cooling-water level decreases, the river sound begins to lag and the river sounds slower. This sound is slightly louder than the rest.

*T+1:15.* As the boiler-hull temperature begins to rise, the sizzling sounds increase. This sound is also slightly louder and precedes the rest.

*T+2:10.* The decreasing cooling water results in a new sample—a babbling brook to be swapped in that significantly lags behind the others.

*T+2:35.* Sizzling increases. The bubbling, popping sound for the fuel water decreases, the bursts of steam and engine run faster, and the crackling fire of fuel oil increases. The above- and below-threshold sounds are beginning to group together.

*T+2:55.* Fuel water temporarily stabilizes (more bubbles), the fans begin to run faster, and the fuel oil stabilizes.

*T+5:15.* The cooling-water brook is barely running and is preceded by an alarm to note a dangerous state. Hull temperature is above threshold as the sizzling increases.

*T+5:20.* Fuel-water and turbine RPMs are below and above threshold, respectively.

*T+5:25.* The sizzling sound and the large-engine sound are preceded by alarms as well as the decreasing bubbling sound.

*T+5:30.* The fuel oil is below threshold as a small crackling fire.

*T+8:30.* As more steam is vented, the fans turn faster and the steam bursts increase frequency.

## Conclusion

*Michael C. Albers*

These techniques represent three ways to use sound in the user interface. Each brings a particular set of strengths to attack the problems of monitoring and troubleshooting complex engineering systems. Furthermore, each technique calls for a different design—a distinct way of applying the technique to the problem and data of interest. Like graphical user interfaces, auditory user interfaces must have a synergy between form and content.

However, even this simplistic example of a complex engineering system clearly exposes the limitations of each individual technique and design. This is not a problem with the designs created here. Rather, each technique brings weaknesses along with its strengths. Sonification is flexible but lacks expressiveness. Earcons are very scalable but require some learning time. Auditory icons are intuitive but are limited in scalability. The “right” answer is rarely to use only one of these techniques without regard for the others.

Once again, thez creation of graphical user interfaces is the same as the creation of auditory user interfaces—use the tools you have on hand to best address the problem you need to solve. By combining certain aspects of each auditory technique and creating a new design, designers can develop a more robust answer to many problems.[17] In fact, the combination of auditory and graphical user interfaces might be the best way to leverage users’ multimodal capabilities when considering any task where people provide information to and gather information from computers.

## References


1. C.D. Wickens, D.L. Sandry, and M. Vidulich, “Compatibility and Resource Competition between Modalities of Input, Central Processing, and Output,” *Human Factors*, Vol. 25, No. 2, June 1983, pp. 227–248.
2. D.D. Woods and E.M. Roth, “Cognitive Systems Engineering,” in *Handbook of Human–Computer Interaction*, M. Helander, ed., Elsevier Science Publishers, Amsterdam, 1988, pp. 3–43.
3. S. Baron, “A Control Theoretic Approach to Modeling Human Supervisory Control of Dynamic Systems,” in *Advances in Man-Machine Systems Research,* Vol. 1, W.B. Rouse, ed., JAI Press, Greenwich, Conn., 1984.
4. V. Vasandani and T. Govindaraj, “Integration of Interactive Interfaces with Intelligent Tutoring Systems: An Implementation,” *Machine Mediated Learning,* Vol. 4, No. 4, Oct. 1994, pp. 295–333.
5. W.W. Gaver, R.B. Smith, and T. O’Shea, “Effective Sounds in Complex Systems: The ARKola Simulation,” *Proc. Computer-Human Interaction Conf. (CHI ’91),* ACM Press, New York, 1991, pp. 85–90.
6. C. Scarletti, “Sound Synthesis Algorithms for Auditory Data Representation,” *Auditory Display: Sonification, Audification, and Auditory Interfaces,* Addison-Wesley, Reading, Mass., 1993, pp. 223–251.
7. S. Barrass, “EarBenders: Using Stories about Listening to Design Auditory Interfaces,” *Proc. First Asia-Pacific Conf. Human-Computer Interaction (APCHI ’96),* Information Technology Inst., Singapore, 1996, pp. 525–538.
8. S. Barrass, “TaDa! Demonstrations of Auditory Information Design,” *Proc. Third Int’l Conf. Auditory Display—ICAD ’96,* Xerox PARC, Palo Alto, Calif., 1996.
9. S. Barrass, “Some Golden Rules for Designing Auditory Displays,” to be published in *Csound Textbook*, B. Vercoe and R. Boulanger, eds., MIT Press, Cambridge, Mass.
10. A.S. Bregman, *Auditory Scene Analysis,* MIT Press, 1990.
11. M. Blattner, D. Sumikawa, and R. Greenberg, “Earcons and Icons: Their Structure and Common Design Principles,” *Human-Computer Interaction,* Lawrence Erlbaum Assoc., Hillsdale, N.J., Vol. 4, No. 1, 1989, pp. 11–44.
12. S.A. Brewster, P.C. Wright, and A.D.N. Edwards, “An Evaluation of Earcons for Use in Auditory Human-Computer Interfaces,” *Proc. ACM/IFIP Interchi ’93,* ACM Press, Addison-Wesley, 1993, pp. 222–227.
13. S.A. Brewster, P.C. Wright, and A.D.N. Edwards, “Experimentally Derived Guidelines for the Creation of Earcons,” *Adjunct Proc. British Computer Soc. Human-Computer Interaction Conf. (BCS HCI ’95),* Springer-Verlag, London, 1995, pp. 155–159.
14. S.A. Brewster, P.C. Wright, and A.D.N. Edwards, “The Design and Evaluation of an Auditory-Enhanced Scrollbar,” *Proc. ACM CHI ’94,* ACM Press, 1994, pp. 173–179.
16. S.A. Brewster, V.-P. Raty, and A. Kortekangas, “Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy,” *Proc. BCS HCI ’96,* Springer-Verlag, 1996, pp. 169–183.
17. M.C. Albers, “The Varese System, Hybrid Auditory Interfaces, and Satellite-Ground Control: Auditory Icons and Sonification in a Complex, Dynamic System,” *Proc. Second Int’l Conf. Auditory Display—ICAD ’94,* Addison-Wesley, 1995, pp. 3–13.