From signal to substance and back: Insights from environmental sound research to auditory display design.

Brian Gygi

Veterans Affairs Northern California Health Care System
Martinez, CA USA 94553
`bgygi@ebire.org`

Valeriy Shafiro

Communications Disorders and Sciences
Rush University Medical Center,
Chicago, IL USA 60612
`Valeriy_Shafiro@rush.edu`

ABSTRACT

A persistent concern in the field of auditory display design has been how to effectively use environmental sounds, which are naturally occurring familiar non-speech, non-musical sounds. Environmental sounds represent physical events in the everyday world, and thus they have a semantic content that enables learning and recognition. However, unless used appropriately, their functions in auditory displays may cause problems. One of the main considerations in using environmental sounds as auditory icons is how to ensure the identifiability of the sound sources. The identifiability of an auditory icon depends on both the intrinsic acoustic properties of the sound it represents, and on the semantic fit of the sound to its context, i.e., whether the context is one in which the sound naturally occurs or would be unlikely to occur. Relatively recent research has yielded some insights into both of these factors. A second major consideration is how to use the source properties to represent events in the auditory display. This entails parameterizing the environmental sounds so the acoustics will both relate to source properties familiar to the user and convey meaningful new information to the user. Finally, particular considerations come into play when designing auditory displays for special populations, such as hearing impaired listeners who may not have access to all the acoustic information available to a normal hearing listener, or to elderly or other individuals whose cognitive resources may be diminished. Some guidelines for designing displays for these populations will be outlined.

1. INTRODUCTION

Compared to the other major classes of familiar sounds, speech and music, environmental sounds are often considered the least important and have certainly been least studied. This is partially because the main function of environmental sounds is conveying information about the physical sources of sound-producing events, which in modern living is less valued than the communicative content of speech or the aesthetic qualities of music.

Nevertheless, environmental sounds are an important component of our everyday listening experience. One of the main benefits listeners with cochlear implants report is an increased ability to perceive environmental sounds [3]. Although appeals to evolutionary aspects of hearing are difficult to prove, it is hard to argue against the notion that the ability to identify sound sources preceded that of listening to speech or music [4]. It would be a more difficult, dangerous and less meaningful world without the ability to recognize environmental sounds.

For the auditory display designer, environmental sounds are useful precisely because of their representational value. While it is difficult to hear speech and not concentrate on the linguistic message, and it is compelling simply to hear music as music, environmental sounds can convey a variety of messages and take on disparate functions: as a warning signal; representing data auditorally; as an icon for carrying out commands on the computer or notifying the user of changes of state in the computer system; or, as themselves, as part of a virtual scene. This paper will focus on their use as auditory icons in computer interfaces, but some other uses, specifically as alarms, will be discussed in the final section.

However, while there has been a great deal of research and standardization in the use of visual icons in computer interfaces, analogous efforts with respect to auditory icons have been somewhat scattered and not well connected. This paper will attempt to incorporate some of the research into auditory icons with established findings from more basic auditory research to further develop the effective use of environmental sounds in auditory displays.

2. BACKGROUND

The theoretical bases for auditory icons were laid out fairly early on in [5]. In it the author, William Gaver, described the different ways information can be mapped to representations, which he labeled “symbolic”, in which the mappings were essentially arbitrary and had to be learned (such as a siren for a warning sign); nomic, where the meaning depended on the physics of the situation, which in the auditory case that would mean inferring a sound-producing source from the acoustics generated in the event; and in between the two are metaphorical mappings, which make use of similarities between the thing to be represented and the representing system. One example would be to use a descending pitch to denote a falling object. There is a continuum between these levels of representation, and a given sound-meaning mapping can change the level with usage. Nomic mappings are the most common way we listen to sounds in the world, namely in terms of the sources that produce the sounds, which he termed everyday listening, as opposed to musical listening, which is more concerned with more symbolic aspects of the sounds. Identifying more environmental sounds is a process of recovering the nomic mappings.

Gaver suggested that auditory displays could be more powerful and useful by taking advantage of everyday listening. The power of nomic mappings comes from the fact that they can convey more multi-dimensional data than simple symbolic associations because of their spectral-temporal complexity. When recognizing an environmental sound, we not only can tell with a great deal of specificity, what objects interacting in what way caused the sound (such as wooden hammer striking a plate) but we can also infer properties of the objects, such as the shape of a struck plate [7], the length of a falling rod [8], and in the case of footsteps, the gender of the walker [9]. Since we are sensitive to change in source properties, Gaver proposed that dimensional information can be conveyed in auditory displays by manipulating a sound in terms of its source properties. For example, if the sound of something dropping into a mailbox is used to inform the use of incoming email, the loudness of the sound can be used to indicate the size of the incoming mail (the louder the sound, the larger the incoming mail) and the timbre of the sound could indicate something of type of mail – if the sound was like a crackling paper, it would indicate a text file.

Gaver pointed out a concurrent property of nomic mappings in auditory displays: they are easier to learn and retain than arbitrary mappings, because they rely on highly learned associations, acquired through our life experience. This was demonstrated in [10]. They used different types of relations in assessing learnability of auditory icons, and found that direct relations, which included both ecological (their term for nomic) and metaphorical relations, were much more learnable than random ones, i.e. symbolic ones, but somewhat surprisingly there was no difference between ecological and metaphorical relations in learnability.

Although the use of environmental sounds was given a solid theoretical foundation, the practical applications were slower to develop, for several reasons. Gaver designed an auditory accompaniment to the Finder feature in Macs, called SonicFinder[6], which had a nicely thought-out interface using a variant of auditory icons in useful ways (a portion of which is detailed in Table 1). For instance, selecting a file was mapped to the sound of an object being tapped, with the type of object indicated by the object material, and size of file represented by the size of the struck object. Although the SonicFinder was often cited as a good example of a sound-based interface, the project was never implemented by Apple.

Finder Events	Auditory Icons
Objects	
Selection	Hitting Sound
Type (file, application, folder, disk, trash	Sound source (wood, metal, etc.)
Size	Frequency
Opening	Whooshing sound
Size of opened object	Frequency
Dragging	Scraping sound
Size	Frequency
Where (windows or desk)	Sound type (bandwidth)
….	….
Windows	
Selection	Clink
Dragging	Scraping
Growing	Clink
Window size	Frequency

Table 1. A partial listing of events in the SonicFinder interface and associated auditory icons adapted from [6].

The end result is that although most operating systems use environmental sounds for a few specialized purposes, such as deleting an object generates a crunching metallic sound, the full functionality of auditory interfaces has not been generally implemented, despite the potential to have a profound impact (e.g. facilitating computer use for the blind). There are several reasons for this, some commercial (the SonicFinder used too much memory in the days of limited memory), but some are due to disadvantages inherent in auditory interfaces, which have been well documented [11, 12]. Sound is not well suited for representing absolute data (as opposed to relative values), the spatial resolution for sounds is not as fine-grained as it is for vision, simultaneous sounds occlude and mask each other more than visual objects, and sound is not a static medium, it unfolds in time, so it cannot represent data instantaneously and thus is problematic for continuous data displays.

However, many of hindrances to full usability of auditory icons come from the relative lack of research into environmental sounds, compared to vision or even speech and music. Although some basic principles for using sounds based on auditory research were outlined in [13], and [14] developed some good heuristics for auditory icon usage, basic research in environmental sounds has so far not been much applied to auditory display design. This paper will discuss some major issues involved in using environmental sounds in auditory displays, and how some recent knowledge gained from the hearing sciences can help to resolve some of these issues.

One major issue is identifiability of the sounds used, that is, how can the designer ensure a sound is actually recognized as intended. The identifiability of the sound depends on both the acoustic properties of the sound, but also, when a sound is presented in an auditory scene, the relationship of the sound to that scene: does the sound “fit” in a semantic sense? Thirdly, what are the best usages of environmental sounds? Which sounds are well-suited for which types of functions? In order to fully utilize the informative capabilities of environmental sounds, it is necessary to enable them to portray changes in states, as [15] suggested. This requires parameterizing the sounds so that certain aspects of the acoustics of the sounds will change, reflecting events in the interface. Finally, designing auditory displays for special populations, such as hearing impaired listeners who may not have access to all the acoustic information available to a normal hearing listener, or to elderly or other individuals whose cognitive resources may be diminished, poses some special considerations.

3. IDENTIFIABILITY

3.1. Baseline Identifiability in the Clear

The identifiability of the environmental sounds used as auditory icons is a major component in their successful use, as noted by [11, 14, 16-18]. Part of that is due to the automaticity of sound recognition. Identifying the source of a sound is one of the basic functions of the auditory system [4] which may also involve a series of cognitive inferences following identification [17], which give the auditory icon its significance. If the sound is unable to be identified, then the usefulness of the icon is greatly diminished[14].

Although humans can identify an immense number of environmental sounds quickly and accurately, studies attempting to determine the baseline identifiability of environmental sounds have yielded mixed results. Often these differences are due to methodological differences. Several early studies [19-22] were hampered by poor quality of sound reproduction or lack of controlled testing conditions. Even among more recent experiments there are differences in the procedures (method of constant stimulus, method of limits), presentation (in the clear, in noise, filtered), and limits on the durations of the sounds presented [14, 23-28]. So, not surprisingly, Ballas, in [17] compared five studies and found quite variable results. However, in [1] identification data were obtained for a group of 52 sounds which had been previously used in three studies of a large corpus of environmental sounds identification [25, 26, 28]. Multiple (3-5) tokens of each sound were tested, for a total of 195 different tokens, and the stimuli were all presented in the clear at an audible level (75 dB SPL) with little or no filtering or editing for length. Listeners identified the sounds using three-letter codes from a list of 90 possible codes. In addition the listeners rated the typicality of the sounds, that is, how good an example of the sound the token they heard was. The results for the best token of each sound, based on 75 young normal hearing listeners (YNH), are shown in Table 2. Forty five of the 52 tokens listed were all recognized at p(c) = 0.9 or greater. In addition, these tokens all had a mean typicality rating of 4 or better on a scale of 1-7 where 1=not typical at all. In addition, the mean p(c) for all tokens for a particular sound is listed. When there is a large discrepancy between the two it means that there were some tokens for a sound that were not well recognized at all.

So there does seem to be a group of sounds that can be recognized reliably in isolation in the clear, as well as a few (e.g., Shovel) that were not well recognized. Further, the correlations between the recognition performance in this test and the data for the same tokens from [26, 28, 29] were quite high, r = 0.88 and 0.83, respectively, indicating that the variance in identification between studies noted in [17] is likely due to the different tokens used in each study.[1]

	Mean p(c)			Mean p(c)	
Label	Token	Sound	Label	Token	Sound
Baby	1.00	1.00	Ice drop	0.97	0.95
Cough	1.00	0.96	Match	0.97	0.93
Dog	1.00	1.00	Bubbles	0.96	0.91
Drums	1.00	0.97	Car accel.	0.96	0.78
Glass	1.00	0.98	Rooster	0.96	0.96
Gun	1.00	0.98	Gargle	0.96	0.95
Laugh	1.00	0.98	Thunder	0.96	0.91
Phone	1.00	1.00	Crash	0.96	0.87
Siren	1.00	0.97	Bells	0.95	0.92
Sneeze	1.00	0.94	Rain	0.94	0.87
Toilet	1.00	0.82	Scissor	0.94	0.81
Whistle	1.00	0.98	B-ball	0.93	0.86
Door	0.99	0.99	Train	0.93	0.88
Clock	0.99	0.93	Claps	0.93	0.80
Helicopter	0.99	0.91	Sheep	0.93	0.89
Gallop	0.99	0.95	Crickets	0.92	0.86
Zipper	0.99	0.99	Waves	0.91	0.80
Cat	0.99	0.88	Pour	0.91	0.87
Cow	0.99	0.98	Tennis	0.90	0.84
Ping-pong	0.99	0.96	Splash	0.89	0.81
Neigh	0.99	0.98	Footstep	0.88	0.82
Bowling	0.99	0.96	Stapler	0.85	0.78
Bird	0.99	0.84	Hammer	0.85	0.62
Typewriter	0.99	0.72	Harp	0.81	0.81
Airplane	0.97	0.83	Cymbal	0.79	0.76
Car start	0.97	0.97	Shovel	0.65	0.53

Table 2. List of sounds tested in [1], the mean p(c) for the best token of that sound, and for all the tokens of that sound as a group.

3.2. Sound Duration and Identifiability

As noted above, sounds unfold in time (hence the theme of this conference) and so the time required to identify an environmental sound needs to be considered in the design of the auditory display. In [14] the author recommended using for auditory icons sounds that were short, with " a wide bandwidth and where length, intensity and sound quality are roughly equal." While that might make sense from a programmatic standpoint, in practice it might lead to choosing icons that all sound like short noise bursts and are thus not easily discriminable. Since environmental sounds represent events in the real world, they will by necessity have different lengths. As mentioned, most of the sounds listed in the previous section were presented with minimal editing, so they were fairly long, with a mean duration of 2240 ms (max 4812 ms, min 224 ms). The correlation between duration and p(c) was essentially zero,

---

[1] Unfortunately, most of the sounds tested were taken from commercial sound effects CDs, which meant they were copyrighted and are not freely distributable. The authors are currently involved in a project of collecting and norming freely distributable high-quality environmental sounds for research purposes.

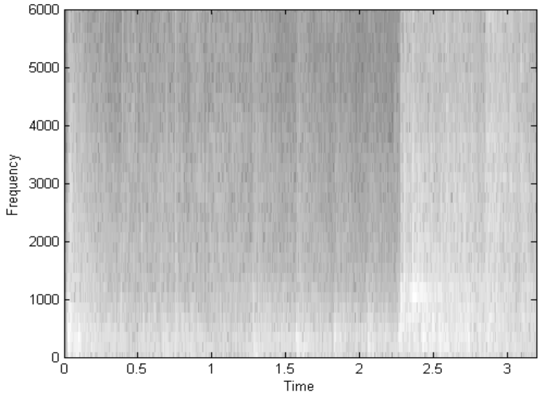


Figure 1 *Bowling ball rolling down a lane and striking pins. The strike occurs 2.3 s into the sound.*

so the shorter sounds were as well recognized as the longer sounds. However, this does not mean that sounds can be edited for length with impunity. In a seminal study of environmental sounds perception, [23], the sounds were all edited to be no more than 625 ms. While some of the sounds preserved their identifiability in this way, for a large number of them the identification accuracy was essentially at chance.

So how much of an environmental sound is necessary to be heard in order to identify it? A study of 117 environmental sounds using a gated paradigm [30] in which the length of the sounds were increased on each presentation by 50 ms found that half of the sounds were identified within 150 ms, which is an amazingly brief span of time. However, the gating paradigm allows listeners multiple exposures to a sound, each time gaining more information about the sound, so the results do not necessarily indicate how listeners would perform at the on the first presentation.

Certainly there are some sounds whose identity is immediately evident. Harmonic sounds can be identified on the basis of their steady-state spectra alone, even when the temporal information is misleading [31]. These include whistles, phones, sirens, musical instruments and animal vocalizations, all of which were among the quickly-recognized stimuli in [30]. So these sounds can be edited for length and still be recognizable. However, many short impulsive sounds tend to have similar envelopes and bandwidths (such as a basketball bouncing, gun, hand claps, door knock, and chopping wood) and the only way to distinguish is in their temporal structure, which includes the periodicity and damping (which was demonstrated in the case of bouncing and breaking bottles in [32]) so a longer sample is necessary in those cases.

Some environmental sounds are actual composite events made up of simpler, more basic events (see [33] for a taxonomy of basic and complex events, discussed below), and in these cases it is necessary to perceive all or nearly of the constituent events to accurately recognize the sound. In an extreme case, the spectrogram of a bowling ball rolling down an alley is plotted in Figure 1. The bowling ball does not strike the pins until 2.3 seconds after the start of the sound. Although this token had a very high identifiability, if it was edited to exclude collision with the pins it would be almost unidentifiable.

3.3. Effects of Filtering on Identification

There may be times when the auditory display designer will wish to filter the sounds used in auditory icons, to avoid masking other important sounds in the display (such as music), or to accommodate a narrow transmission bandwidth, or for special populations who may be using the display. For example [34] described an auditory GUI in which auditory icons were high- or lowpass filtered to indicate whether they were highlighted or deselected, respectively (see section 6 on auditory icons for special populations). Since environmental sounds are a result of the whole catalog of possible sound-producing events in the world, there is an extreme of spectral-temporal variation among environmental sounds, with some sounds with a strong harmonicity (vocalizations), others more like broad band noises (water sounds), some are relatively steady state (air sounds) and some are transient (impact sounds).

As a result, the effects of filtering on environmental sounds are not uniform, as was demonstrated in [25]. Figure 2 is adapted from that piece and shows the effects of low- and high-pass filtering on a selected group of environmental sounds. There are some sounds, such as weather sounds, like thunder

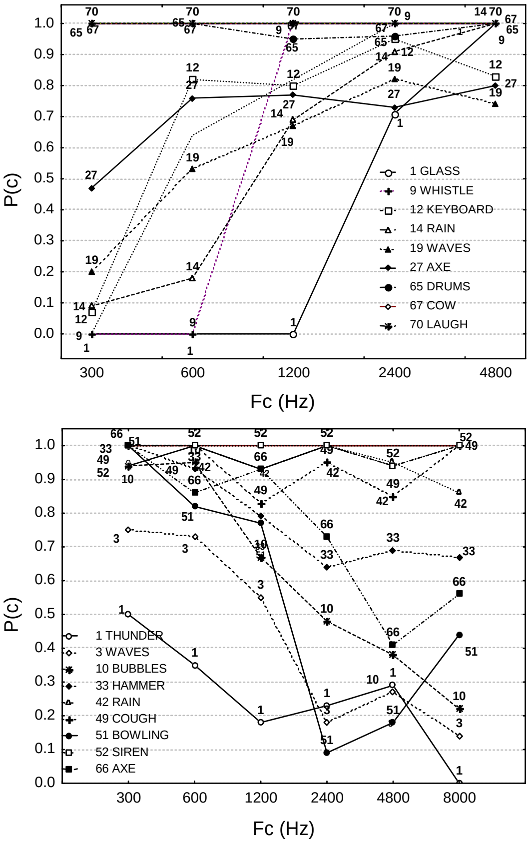


Figure 2. *Effects of lowpass (top) and highpass (bottom) filtering on a group of environmental sounds reprinted from [25] with the authors' permission. The cutoff frequency for the filters are shown on the abscissa.*

and waves that are extremely adversely affected by low-pass filtering but are resistant to high-pass filtering. Some sounds, like an axe chopping, carry all the information in the envelope and so are robust to most filtering.

3.4. Effects of Context on Identification of Environmental Sounds

Sounds in the world do not happen in isolation, but rather occur as part of an auditory scene concurrently with other related events. In any given setting there are sounds that are more or less likely to occur, i.e. that are congruent or incongruent with that setting. Although context has been found to be helpful in perceiving speech sounds [35] and in various psychophysical tasks [36-38], the effects of context on identification of environmental sounds have not been extensively researched.

It is often assumed that it is more difficult to recognize sounds out of context (or in no context at all) [14]. However [39] tested the identifiability of pairs of confusable sounds which were embedded in a sequence of other sounds that provided a scenario, such as a match lighting, a fuse burning, and an explosion. The sequences were designed to be either consistent with the true source of the sound, biased towards the other member of the test sounds pair (e.g., slicing food, chopping food and fuse burning), or random.

The effect of consistent context significantly raised performance above that of the biased sequences and the random sequences; however it did not improve performance over a baseline condition. The authors interpreted this finding as showing that "…the only positive effect of consistent contexts is to offset the negative effects of embedding a sound in a series of other sounds."

The effects of embedding environmental sounds in more naturalistic auditory scenes (as opposed to a series of isolated sounds) was tested in [40] using field recordings of scenes as backgrounds and as targets some of the environmental sounds which were used in other studies (e.g., [1]). The sound-scene combinations were designed to be either congruent (the target sound was considered likely to appear in a scene, such as a horse in a barnyard), incongruent (a horse in a restaurant) or neutral (which were used as foils). Figure 3 shows the identification results for highly trained listeners at Sound-Scene ratios (So/Sc) of -18, -15 and -12 dB. Somewhat surprisingly, the incongruent sounds were more easily identified than the congruent ones, except at the very poor So/Sc ratio of -18, when there was no significant difference. It appears from these data that context, at least the naturalistic contexts used here, did not enable identification of expected sounds, but rather of unexpected sounds.

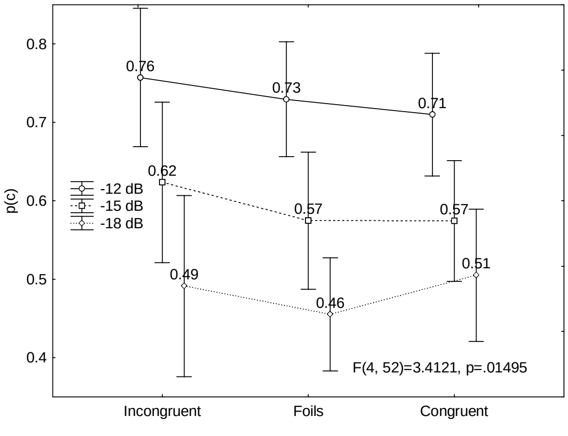


Figure 3 *Identifiability of environmental sounds in incongruent, congruent, or neutral contexts as a function of presentation level. From [40] with the author's permission.*

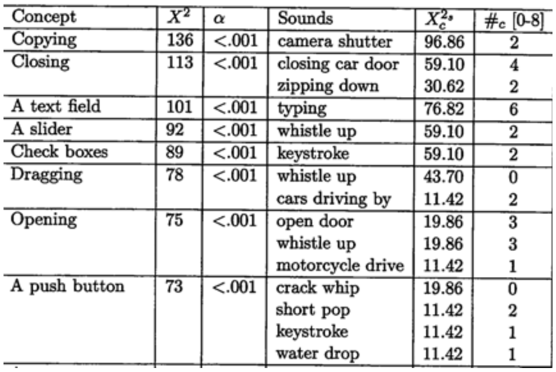
The authors interpreted the results as reflecting the tendency of the auditory system to detect change or novel stimuli, which is certainly adaptive from an evolutionary standpoint and not unique to hearing, but has been found in vision as well e.g., [41]. So we tend to group expected sounds into a background which do not compel attention, so it takes a greater saliency (in this case a louder stimulus) to stand out. In contrast, we tend to monitor an auditory scene for unexpected events, so as to be better able to react to them. Designers of auditory displays can utilize this tendency to either enhance the saliency of an auditory icon or to make it less noticeable, depending on the need. However, the fact that this finding is in direct contrast to the effects of context in speech stimuli indicates that it may be only applicable to naturalistic environments, which are based on probabilities of occurrence and does not hold for rule-based systems, such as grammar, in which there are stronger constraints on what sequences are possible. Since auditory displays tend to fall between the two extremes, i.e., not anything is possible but they are still not entirely deterministic, it is possible that listeners might switch between the modes of attending given the setting and task demands.

4. SEMANTIC COMPATIBILITY OF SOUND AND FUNCTION

When a sound is identified, as noted in [17] there are numerous associations activated in the listener's pertaining to knowledge of the source: how big the objects involved in the source are, what they are made of, how they were interacting. In order to use them effectively in a display, the function in the display should relate in some way to the source properties already known by the user [42]. However, the number of possible mappings from physical sources to auditory icons is limited, and effective ones may not always be obvious or intuitive.

In the SonicFinder interface, Gaver developed a complex mapping of sounds to functions, some of which were listed above in Table 1. While some are obvious and have been incorporated into modern interfaces (the sound of dropping something in a trash can to denote deleting it), others are less intuitive and on the "symbolic" end of the scale, e.g., a pouring sound for copying. As a result, these mappings would likely be harder for the listener to learn.

There are still no standards for which sounds are recommended to use with which computer-generated events, so most auditory designers have to go by their own intuitions. In [14] the author conducted a study to determine which auditory icons eight subjects felt mapped best to which interface concepts (her term for computer events). A partial listing of the results is in Table 3. Many of the results were expected (to close a file the subjects preferred the sound of a door closing), but several were somewhat surprising: for instance, the overwhelming preference to indicate copying was the sound of a camera shutter. Although her work provided several good



Concept	X^2	α	Sounds	X_c^{2*}	$\#_c$ [0-8]
Copying	136	<.001	camera shutter	96.86	2
Closing	113	<.001	closing car door	59.10	4
			zipping down	30.62	2
A text field	101	<.001	typing	76.82	6
A slider	92	<.001	whistle up	59.10	2
Check boxes	89	<.001	keystroke	59.10	2
Dragging	78	<.001	whistle up	43.70	0
			cars driving by	11.42	2
Opening	75	<.001	open door	19.86	3
			whistle up	19.86	3
			motorcycle drive	11.42	1
A push button	73	<.001	crack whip	19.86	0
			short pop	11.42	2
			keystroke	11.42	1
			water drop	11.42	1

Table 3. A partial listing of auditory interface concepts and users' preferred associated auditory icon adapted from [2].

guidelines it has not been adopted as a standard, but provides a good starting point. Similarly, [43] obtained blind and sighted users' ratings of the recognizability and appropriateness of auditory icons, the results of which are discussed in Section 6.

Until more research with a wider range of sounds and computer events is conducted (or standards are adopted) display do designers just have to rely on guesswork and what appeals to them when associating a sound with an unconventional event. Since environmental sounds activate a range of semantic features, if one knew what those features are, it might point the way to a good mapping. The perceptual dimensions of 145 environmental sounds were measured in [44] using the semantic differential method, in which listeners rated sounds on twenty different scales based on binary pairs of qualities, such as "tense – relaxed", "light – heavy" or "compact – scattered". The ratings can then be subjected to factor analysis to determine the primary underlying perceptual dimensions. The authors found four main perceptual dimensions: harshness, complexity, appeal and size. So if a sound designer has an activity that might seem tense, or complex to the user, (s)he can use a sound that rated highly on those scales (the loading of each sound on each factor is supplied in the article).

5. USING ENVIRONMENTAL SOUNDS TO REPRESENT CONTINUOUS DATA

In the SonicFinder interface, the sound of something dropping in a basket denoted incoming mail, and the larger the incoming mail, the louder the sound. This is an example of how environmental sounds can represent not just events but a range of values for an event. However, to do this requires parameterizing the acoustics of the sounds to match the range of values of data displayed. This is a nontrivial problem. While certain salient acoustic variables such as loudness are quite easily manipulated, others (such as spectral centroid,) are not. Further, as was pointed out in [45], for auditory icons to be consistent with the nomic relations of the underlying sound, the parameterization should not just be affected in an arbitrary acoustic feature, but must reflect source properties of the sound. So, in the example cited above, the louder sound was an effective indicator of the size of an incoming mail message because larger things do tend to make louder sounds.

Unfortunately, the acoustical effects of variations in sound-producing events are seldom so neat. If one wanted to indicate an approaching or receding object using sound, one could also make the sound louder or softer, but that would only be a crude approximation of the actual effect because there are also significant spectral changes when objects approach or recede [46] – high frequencies tend to be more greatly attenuated than low frequencies with distance. For sounds that are produced by the interactions of two objects, such as impact sounds, the acoustics are affected by the materials of the interacting objects, the shapes of the interacting objects and of course the type of interaction[7, 9, 47-49]. So to be able to adequately manipulate the sounds to approximate changes in source properties, fairly detailed physical models of the sounds are needed. Although standard equations exist in acoustic textbooks for a number of interactions, Gaver [15, 50] pointed out that the parameters of the models should reflect physical source attributes, which are realized in the acoustic domain, as shown in Table 4.

Since the goal of everyday listening is sound source recognition, Gaver developed a taxonomy of basic sound-producing events [33], which he described as impacts, liquid sounds and aerodynamic sounds and looked at physics of the events to develop models. Gaver implemented this in a model of a class of impact sounds (striking an object with a mallet) with parameters that could be constrained to patterns typical of various object configurations, material, and mallet hardness. He tested sounds generated using these on four listeners who confirmed that the impact sounds generated in this way were realistic [50]. Gaver also proposed models for liquid sounds, scraping sounds and machine sounds, although he said that none of the other models were as successful and he did not report any quantitative tests of them.

Gaver's approach has been very influential and has lead to numerous physical models for environmental sounds based on source properties, although more recent work involving similarity judgments [51] has indicated that listeners tend to regard liquids and aerodynamic sounds as a single class of sound-producing events, and consider harmonic sounds as a separate class of sounds. Most of these models have been for impact sounds, e.g. [52-55], but there have also been models of rolling sounds [56, 57], liquid sounds [58] and rubbing sounds [59]. For more information, the website for the Sounding Objects project (SOb), www.soundobject.org, has a large repository of papers and reports in this area.

The numerosity of these models has made a comprehensive evaluation difficult, but it is important to

Frequency Domain	*Temporal Domain*
Restoring force (material feature)	Interaction type
Density (material feature)	Damping (material feature)
Size (configurational feature)	Internal Structure (material feature)
Shape (configurational feature)	Support (configurational feature)
Support (configurational feature)	

Table 4. Physical parameters of an impact event and their acoustic manifestations, either in the frequency to the time domain. Adapted from [50].

remember that there may be no “best” way to synthesize sounds parametrically. As long as the pertinent information is available and audible, listeners are remarkably good at attending to the relevant features for extracting that information [60]. In fact [61] showed that due to the redundancy of spectral-temporal information in environmental sounds even listeners who do not adopt the most efficient strategy for identifying a sound still manage to perform nearly as well as an ideal listener. The sensitivity of listeners to minute changes in environmental sounds has received very little attention: one study showed that thresholds for discriminating changes in specific spectral regions of a group of environmental sounds were relatively large, on the order of 7-10 dB, although there were large individual differences [62]. In addition, it is not necessarily the case that the most “realistic” sound is the best. Cartoonification is as useful for auditory icons as it is for visual icons in terms of memory storage and interface continuity [63]. There are even instances where “fake” sounds are judged to be better than the “real thing” [64]. So when assessing physical models, ease of implementation should be a consideration. Along those lines it should be noted that the website for Dick Hermes’ Sound Perception class has downloadable Matlab code home.tm.tue.nl/dhermes/lectures/sd/ChV.html for generating several different type of sound-producing events, such as impacts, bouncing sounds, rolling sounds, dripping and machine sounds.

Alternatively, if a designer really wants to use actual sounds and have some variation in the source properties, the website for the Auditory Perception Laboratory http://titan.cog.brown.edu:8080/AuditoryLab/ has numerous high-quality recordings of actual sound-producing events such as rolling, deformation, liquid, air and impact events, with a variety of materials and type of interacting objects. For example, the rolling sounds have a variety of accelerations so that a parameterizing could be simulated (although the memory requirements would be quite substantial).

6. AUDITORY ICONS FOR SPECIAL POPULATIONS

Most of the previous research has been conducted on subjects with normal hearing and cognitive abilities. However, given that a great number of Web applications are for special populations, it is useful to know what considerations need to be taken into account regarding use of environmental sounds for these populations.

6.1. Hearing-Impaired Users

An obvious group that must be considered are hearing impaired users, those using hearing aids (HA), or cochlear implants (CI), as well as those without any prosthetic hearing device at all, especially in the area of alerts or warning sounds. Although all of these groups have greatly diminished auditory sensitivity, the nature of the information they are able to perceive varies drastically. Hearing impaired persons most often have varying losses of sensitivity in the upper frequencies (typically > 2kHz) although hearing at 1000 Hz and below may be near normal. This effect can be approximated by having a lowpass filter with a cutoff at somewhere between 2 and 4 kHz, and the sounds that were found in [25] to be identifiable with only low-frequency energy preserved (see Figure 2) will be best-suited for hearing-impaired people who are not using any amplification device.

For hearing aid users the situation is somewhat different. Although hearing aids adjust the gain on those frequencies the user have trouble hearing to compensate for these losses, many hearing aids users report extreme sensitivity to loudness (recruitment) particularly in those regions and as a result the dynamic range between what is audible and what is painful is drastically reduced, from 100 dB to 60 dB or less [65]. So when using environmental sounds as auditory icons, care should be taken that the overall level of the icons does not greatly exceed the level of the other sounds in the display and that sounds with a large amount of high frequency content are avoided. It should also be noted that a comparison of responses on an open – ended questionnaire of individuals with hearing loss with those of experienced hearing aid users revealed that the number of problems with environmental sound perception was comparable between the two groups [66].

Cochlear implants users pose quite different challenges for the auditory display designers. Cochlear implants are given to profoundly deafened persons, who have little or no residual hearing. The processing in cochlear implants involves filtering using a bank of bandpass filters, extracting the envelope from each of the filter bands, and using each envelope modulate the amplitude of electrical pulses delivered by implant electrodes. The results can be simulated by using the envelopes to modulate the amplitude of a particular carrier signal (typically a noise band or a sine wave) specific to the frequency of each band. A schematic is shown in Figure 4 with a white noise carrier and a six-channel bandpass filterbank to produce a six-channel Event-Modulated Noise (EMN). The result is the temporal information of the original signal is largely preserved, while the spectral information is greatly reduced.

What this means is that CI users have access to temporal cues but much less to spectral cues. While this is adequate for speech perception purposes (at least in the quiet –listeners can achieve near-perfection speech recognition scores with as few

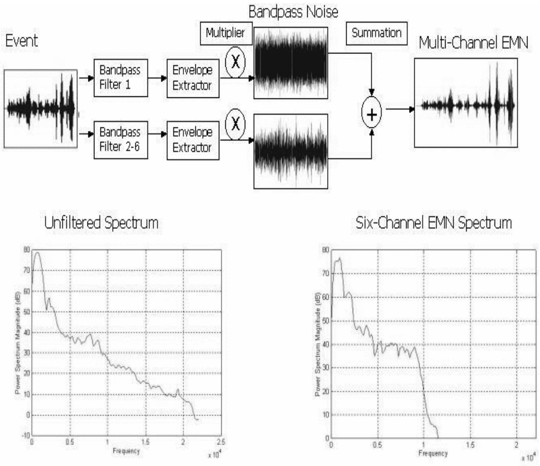


Figure 4. *A schematic of the processing involved in creating Event Modulated Noises (EMF) which simulate the effects of cochelar implants on environmental sounds.*

as four channels [67]) the effect on perception of environmental sounds is more complex. Studies testing environmental sounds identification using CI simulations on normal-hearing listeners [25, 28] and in actual CI users [29, 68, 69] have both indicated that some environmental sounds are perceived quite accurately using CI processing, while others are almost unidentifiable.

The main factor determining this seems to be the amount of temporal versus spectral information in the target sound. Sounds which have a distinct envelope but whose spectra are similar to a broadband noise, such as a horse running, hammering or a ping-pong ball bouncing, are perceived quite accurately. In contrast, sounds with a steady-state envelope, such as a car horn, wind blowing or a flute, tended not to be recognized well at all. One mitigating factor is the number of channels used in the processing scheme, since the greater number of channels, the more spectral information is available to listeners. So sounds such as a baby crying or a bird chirping, which have harmonic spectra, but also a unique temporal patterning, can be quite well identified with a greater number (4-8) of channels. [28] also showed that the number of frequency channels required for 70% correct identification (i.e., 8 channels or less or 16 channels or more) can be predicted with relatively high accuracy of 83% based on the number of bursts and the standard deviation of centroid velocity in the orginal environmental sound.

It should be noted that overall performance on identifying environmental sounds actually may decrease with greater than 16 channel processing [28]. However, the designer of auditory displays for CI users has no knowledge of, or any way of knowing, the number of channels in the CI users' processing scheme, so the best strategy may just be assume the processing scheme that provided the best overall performance in [28] (8 channels) and use sounds that were well identified in that condition (the appendix of [28] has a handy listing of the sounds that were identified with at least 70% accuracy in each channel condition).

One issue that is common to all the hearing impaired populations is a decrease in the ability to separate sounds in a mixture of sounds [70-72], even among hearing impaired listeners fitted with hearing aids. This failure of the "cocktail party effect" [73] means that auditory displays designed for hearing impaired users cannot contain too many concurrent sounds as the users will likely not be able to hear them, which means that many of the context effects mentioned earlier will be difficult to utilize (i.e. it will be more difficult to make an incongruous sound stand out in hearing impaired users).

6.2. Cognitive Factors in Identifiability

Although peripheral processing is a major determinant of identifiability, there are also cognitive factors which can cause hearing impairment. One of these is the general cognitive slowing which occurs with aging [74], which has a strong effect on speech perception: [75] found that declines in speech perception in elderly listeners went above that expected from purely audiometric factors.

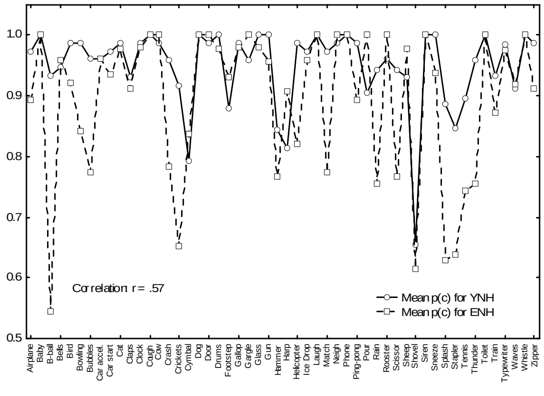


Figure 5. *Mean p(c) for young and elderly listeners for the sounds tested in [1]*

In terms of environmental sound perception, the study mentioned earlier which tested baseline identification of a large group of environmental sounds in young normal listeners (YNH) [1] also tested 50 elderly listeners with normal audiograms, adjusted for age (ENH). The mean identification accuracy for the ENH was significantly less than for the YNH, 0.89 vs. 0.95, even though the same tokens were presented in the clear to both groups. Figure plots the mean performance on the different sounds for both. Other than the elderly listeners being overall worse at identifying, there are some notable differences on particular sounds: for example, elderly listeners were quite bad at identifying a basketball bouncing, which the young listeners were nearly perfect at. There were also some interesting discrepancies in which particular token was judged to be the "most typical" token for a certain sound: among the airplane tokens, YNH judged a recording of a jet to be the most typical of airplane sounds, whereas ENH thought a prop plane was.

As with the hearing impaired, normal-hearing elderly also have greater difficulty with isolating a sound in a mixture of sounds, which is true both for speech [76] and environmental sounds. The study testing target sounds in natural backgrounds described earlier [40] was replicated with ENH listeners. In addition to overall poorer identification of sounds in all contexts, the Incongruency Advantage described earlier for YNH only occurs in elderly listeners at about 6 dB greater So/Sc levels.

6.3. Visually-Impaired Users

One of the main groups that can benefit from improved auditory displays are visually-impaired users, who must get the vast majority of their information through the auditory modality. [43] showed that using auditory icons significantly speeded up blind users' performance on a word processing task. A large number of Web browsers for the visually-impaired have merely translated the browser into text using a screen reader which is read to the user. There are some applications that go further and try to represent some of the embedded images and functions aurally, such as WebbIE and eGuideDog. However, few commercial applications have taken advantage of the work that has been on developing richer auditory interfaces for the blind.

One major problem with translating Web pages for the blind is converting the dense visual information which is presented in two dimensions to an auditory presentation which is time-bound. It is possible to map visual coordinates to binaural using Head-Related Transfer Functions (HRTFs) but as noted in the Introduction, spatial resolution for hearing is much poorer than

Interface Object	*Sound*
Editable text area	Typewriter, multiple keystrokes
Read-only text area	Printer printing out a line
Push button	Keypress (ca-chunk)
Toggle button	Pull chain light switch
Radio button	Pa pop sound
Check box	One pop sound
Window	Tapping on glass (two taps)
Container	Opening a door
Popup dialog	Spring compressed then extended
Application	Musical sound
Action	*Nonspeech Auditory Feedback*
Selection	Ripping papers
Switching between apps.	Paper shuffling
Navigation error	Ball rebounding against wall
Entering text mode	Rollin or rocking sound (drawer pulled out)
Moving edit cursor in text area	Click &itch based on position in text)
Popup appearing or disappearing	Whistle up or down
Application connecting to Mercator	Winding
Application disconnecting	Flushing
Selection	Ripping papers
Switching between applications	Paper shuffling

Table 5. Auditory icons associated with graphical objects. Adapted from [78].

for vision and in binaural presentations there is nearly always masking between components even if they are quite far apart in the display. Such a system was implemented using auditory icons in [77], but no evaluation data were supplied, nor were the types of auditory icons used described.

A more difficult problem is how to present auditory icon that need to be continually available onscreen because the display will very quickly become cluttered with continually sounding icons., causing the user to be unable to discriminate them and overwhelmed with too much information.

A notable early attempt at designing an auditory-only graphical interface was the Mercator Project [78]. The author carefully examined visual GUIs in term of the structure and function of the objects and translated those to auditory objects .in a hierarchical arrangement (similar to telephone menus) rather than a spatial one. To create auditory icons that resembled visual ones, she tried to find sounds that convey the affordances of the visual icons. A listing of the objects and their associated auditory icons is in Table 5. To parameterize the sounds, as mentioned earlier, filtering was used to “thin” (high pass filter) or muffle (low-pass filter) the sounds to represent highlighting or deselecting of an icon, and to represent properties such as the number of lines (for a text object) of the number of children for a container object. Finally, a number of actions of the system were also associated with auditory icons, which are also listed in Table 5. For example if a user wants to cancel and close a popup window, the sequence would be:

"Rip" - the selection is successful.
"Whistle-down" - the pop-up disappear from the screen.
"Ca-chunk" "Reply" - the user is moved back to the reply push button.

According to the authors, reactions from the blind users trained in Mercator was positive, (and the use of environmental sounds was “overwhelmingly positive”) although no controlled studies were done on the blind. A study conducted with sighted persons showed that although users learned the auditory interface quite quickly, there was little benefit from first becoming familiar with a standard GUI, suggesting that navigating the two system required different skills (not surprising, given the differences between the two).

Although, unlike some of the other system described here, Mercator was developed commercially by Sun, it seems to have gradually disappeared from their supported software and as of 2006 was declared dead. So there is still a need for a functional auditory interface that can truly translate visual graphics into appropriate auditory icons. For the designer who wishes to undertake such a project, there are some factors to keep in mind.

Blind users, can be assumed to have normal hearing and cognitive skills. In terms of sensitivity to environmental sounds, [79] reported no differences between blind and sighted persons in accuracy for identifying environmental sounds (although the overall performance in both groups was rather low, p(c) = 0.76-0.78. Comparing sighted with blind users’ ratings of mapping from auditory icons to interface events, [80] found blind users gave significantly lower overall ratings of the appropriateness of the mappings, perhaps indicating that blind users have stronger associations of sounds to physical events, and so to abstract away from the established nomic relationship to a more metaphorical one is more difficult. This may be alleviated with training, but it does pose a challenge for designing an auditory interface that will be immediately usable for visually-impaired listeners.

7. SUMMARY

This paper has attempted to outline some of the factors auditory display designers need to be mindful of when using environmental sounds for auditory icons:

1) Identifiability. The source of the original sound must be easily and rapidly identifiable for the users. A large number of common environmental sounds can be readily recognized by users, but not all, and the identifiability suffers if there is any filtering or editing of the sounds.

2) Context. If one auditory icons is embedded in a number of other sounds, the context can either enhance identifiability or detract from it, depending on contextual congruency of the target icon sound and other background sounds. An incongruent sound can be more identifiable if listener is monitoring for expected sounds.

3) Mapping (i.e. choosing the right sound for the right icon). Some sounds more readily lend themselves to certain functions in a display (for instance the sound of a closing door when shutting down a program) but these are not standardized. The semantic attributes of a number of environmental sounds have been established and the designer can use these to find an appropriate sound for an icon.

4) Parameterization of the underlying sounds. To represent continuous data with an auditory icon, the underlying sound must be manipulable in terms of the sound source properties.

There are several physical models of environmental sounds that can achieve this; choosing the right one will depend on ease of implementation, memory requirements and degree of fidelity to the original sound desired.

5) Target population. When designing displays for special populations, special considerations must be made if the target population is hearing impaired or has a cognitive impairment. For hearing impaired without hearing aids, sounds that depend on high frequency information for identification should be avoided. For hearing aid users, the dynamic range of the sounds should be compressed to avoid causing pain to the listeners. Cochlear implant recipients will perform best with sounds that rely largely on temporal structuring. Elderly listeners generally perform less well on sound identification than young listeners, even elderly with normal hearing. However, for some sounds they perform similarly to young listeners. All of these groups have difficulty when there are numerous sounds in the display and they have to focus on a single sound. For these users, the display designer should take care than the auditory display does not have too many concurrent sounds.

8. REFERENCES

[1] Gygi, B., "Studying environmental sounds the Watson way", J Acoust Soc Am, vol. 115, no. 5, pp. 2574, 2004

[2] Mynatt, E. "Designing with auditory icons". in *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Santa Fe, NM, U.S., 1994.

[3] Zhao, F., S.D.G. Stephens, S.W. Sim, and R. Meredith, "The use of qualitative questionnaires in patients having and being considered for cochlear implants", Clin Otolaryngol, vol. 22, pp. 254-259., 1997

[4] Warren, R.M., *Auditory perception and speech evolution.*, in *Annals of the New York Academy of Sciences, Origins and Evolution of Language*. 1976. p. 708-717.

[5] Gaver, W.W., "Auditory icons: using sound in computer interfaces", Hum.-Comput. Interact., vol. 2, no. 2, pp. 167-177, 1986

[6] Gaver, W.W., "The SonicFinder: An Interface that Uses Auditory Icons. " Hum.-Comput. Interact., vol. 4, no. 1, pp. 67-94, 1989

[7] Kunkler-Peck, A.J. and M.T. Turvey, "Hearing shape", J Exp Psychol Human, vol. 26, no. 1, pp. 279-294, 2000

[8] Carello, C., K.L. Anderson, and A.J. Kunkler-Peck, "Perception of object length by sound", Psychol Sci., vol. 9, no. 3, pp. 211-214, 1998

[9] Li, X., R.J. Logan, and R.E. Pastore, "Perception of acoustic source characteristics: Walking sounds", J Acoust Soc Am, vol. 90, no. 6, pp. 3036-3049, 1991

[10] Keller, P. and C. Stevens, "Meaning From Environmental Sounds: Types of Signal-Referent Relations and Their Effect on Recognizing Auditory Icons", J Exp Psychol-Appl, vol. 10, no. 1, pp. 3-12, 2004

[11] Absar, R. and C. Guastavino. "Usability of Non-Speech Sounds in User Interfaces". in *Proceedings of the 14th International Conference on Auditory Display*. Paris, France, 2008.

[12] Fernstroem, M., E. Brazil, and L. Bannon, "HCI Design and Interactive Sonification for Fingers and Ears", IEEE MultiMedia, vol. 12, no. 2, pp. 36-44, 2005

[13] Watson, C.S. and G.R. Kidd. "Factors in the design of effective auditory displays". in *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Santa Fe, NM, U.S., 1994.

[14] Mynatt, E. "Designing with auditory icons" in *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Santa Fe, NM, U.S.: 1994 of Conference.

[15] Gaver, W.W., *Using and creating auditory icons*. SFI studies in the sciences of complexity, ed. G. Kramer. Addison Wesley Longman, 1992.

[16] Ballas, J., *Delivery of information through sound*. SFI studies in the sciences of complexity, ed. G. Kramer. Addison Wesley Longman, 1992.

[17] Ballas, J.A. "What is that sound? Some implications for sound design." in *Design Sonore*. Paris, 2002.

[18] Lucas, P.A. "An evaluation of the communicative ability of auditory icons and earcons". in *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Santa Fe, NM, U.S., 1994.

[19] Lass, N.J., S.K. Eastman, W.C. Parrish, S. K.A., and D. Ralph, "Listeners' identification of environmental sounds", Percept Motor Skill, vol. 55, no. 1, pp. 75-78, 1982

[20] Miller, J.D. and D.C. Tanis, "Recognition memory for common sounds", Psychon Sci, vol. 23, no. 4, pp. 307-308, 1973

[21] Lawrence, D.M. and W.P. Banks, "Accuracy of recognition memory for common sounds", Bull Psychonom Soc, vol. 1, no. 5A, pp. 298-300, 1973

[22] Vanderveer, N.J., "Ecological acoustics: Human perception of environmental sounds", Dissertation Abstracts International, vol. 40, no. 9-B, pp. 4543, 1980

[23] Ballas, J.A., "Common factors in the identification of an assortment of brief everyday sounds", J Exp Psychol Human, vol. 19, no. 2, pp. 250-267, 1993

[24] Fabiani, M., V.A. Kazmerski, and Y.M. Cycowicz, "Naming norms for brief environmental sounds: Effects of age and dementia", Psychophysiology, vol. 33, no. 4, pp. 462-475, 1996

[25] Gygi, B., G.R. Kidd, and C.S. Watson, "Spectral-temporal factors in the identification of environmental sounds", J Acoust Soc Am, vol. 115, no. 3, pp. 1252-65, 2004

[26] Marcell, M.M., D. Borella, M. Greene, E. Kerr, and S. Rogers, "Confrontation Naming of Environmental Sounds", J Clin Exp Neuropsyc, vol. 22, no. 6, pp. 830-864, 2000

[27] Myers, L.L., T.R. Letowski, K.S. Abouchacra, J.T. Kalb, and E.C. Haas, "Detection and recognition of octave-band sound effects", J Am Acad Audiol, vol. 7, pp. 346-357, 1996

[28] Shafiro, V., "Identification of environmental sounds with varying spectral resolution." Ear Hear, vol. 29, no. 3, pp. 401 - 420, 2008

[29] Shafiro, V., B. Gygi, M.-Y. Cheng, M. Mulvey, and B. Holmes, "Perception of speech and environmental sounds in cochlear implant patients", J. Acoust. Soc. Am., vol. 123, no. 5, pp. 3303, 2008

[30] Guillaume, A., L. Pellieux, V. Chastres, and C. Blancard. "How long does it take to identify everyday sounds". in *ICAD 04 -Tenth Meeting of the International Conference on Auditory Display*. Sydney, Australia, 2004.

[31] Gygi, B., "From acoustics to perception: How to listen to meaningful sounds in a meaningful way", J. Acoust. Soc. Am., vol. 113, no. 4, pp. 2326, 2003

[32] Warren, W.H. and R.R. Verbrugge, "Auditory perception of breaking and bouncing events: A case study in ecological acoustics", J Exp Psychol Human, vol. 10, no. 5, pp. 704-712, 1984

[33] Gaver, W.W., "What in the world do we hear? An ecological approach to auditory event perception", Ecol Psychol, vol. 5, no. 1, pp. 1-29, 1993

[34] Mynatt, E.D., "Transforming graphical interfaces into auditory interfaces for blind users", Human-Computer Interaction, vol. 12, no. 1, pp. 7-45, 1997

[35] Bilger, R.C., J.M. Nuetzel, W.M. Rabinowitz, and C. Rzeczkowski, "Standardization of a test of speech perception in noise", J Speech Hear Res, vol. 27, pp. 32-48, 1984

[36] Hafter, E.R. and K. Saberi, "A level of stimulus representation model for auditory detection and attention", J. Acoust. Soc. Am., vol. 110, no. 3, pp. 1489, 2001

[37] Schlauch, R.S. and E.R. Hafter, "Listening bandwidths and frequency uncertainty in pure-tone signal detection", J Acoust Soc Am, vol. 90, pp. 1332-1339, 1991

[38] Watson, C.S. and D.C. Foyle, "Central factors in the discrimination and identification of complex sounds", J. Acoust. Soc. Am., vol. 78, no. 1, pp. 375-380, 1985

[39] Ballas, J.A. and T. Mullins, "Effects of context on the identification of everyday sounds", Hum Perform, vol. 4, no. 3, pp. 199-219, 1991

[40] Gygi, B. and V. Shafiro, "The Incongruency Advantage for Environmental Sounds Presented in Natural Auditory Scenes", Submitted, vol., 2008

[41] Gordon, R.D., "Attentional allocation during the perception of scenes. " J Exp Psychol Human, vol. 30, pp. 760-777., 2004

[42] Lucas, P.A. "An evaluation of the communicative ability of auditory icons and earcons" in *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Santa Fe, NM, U.S.: 1994 of Conference.

[43] Petrie, H. and S. Morley. "The use of non-speech sounds in non-visual interfaces to the MS Windows GUI for blind computer users". in *Proceedings of the 5th International Conference on Auditory Display (ICAD98)*. University of Glasgow, U.K., 1998.

[44] Kidd, G.R. and C.S. Watson, "The perceptual dimensionality of environmental sounds", Noise Control Engineering Journal, vol. 51, no. 4, pp. 216-231, 2003

[45] William, W.G., "Auditory icons: using sound in computer interfaces", Hum.-Comput. Interact., vol. 2, no. 2, pp. 167-177, 1986

[46] Coleman, P.D., "An analysis of cues to auditory depth perception in free space", Psychological Bulletin, vol. 60, no. 3, pp. 302, 1963

[47] Freed, D., "Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events." J Acoust Soc Am, vol. 87, no. 1, pp. 311-322, 1990

[48] Grey, J.M., "Multidimensional perceptual scaling of musical timbres", J. Acoust. Soc. Am., vol. 61, no. 5, pp. 1270-1277, 1977

[49] Lakatos, S., S. McAdams, and R. Caussé, "The representation of auditory source characteristics: Simple geometric form", Percept Psychophys, vol. 59, no. 8, pp. 1180-1190, 1997

[50] Gaver, W.W., "How do we hear in the world? Explorations in ecological acoustics", Ecol Psychol, vol. 5, no. 4, pp. 285-313, 1993

[51] Gygi, B., G.R. Kidd, and C.S. Watson, "Similarity and Categorization of Environmental Sounds", J Acoust Soc Am, vol. 69, no. 6, pp. 839-855, 2007

[52] Aramaki, M. and R. Kronland-Martinet, "Analysis-synthesis of impact sounds by real-time dynamic filtering", IEEE T Speech Audi P, vol. 14, no. 2, 2006

[53] Avanzini, F. "Synthesis of Environmental Sounds in Interactive Multimodal Systems". in *Proceedings of the 13th International Conference on Auditory Display (ICAD2007)*. Montreal, Canada, 2007.

[54] Lee, J.-F., I.Y. Shen, J. Crouch, W. Aviles, D. Zeltzer, and N. Durlach, "Using physically based models for collision-sound synthesis in virtual environments", J. Acoust. Soc. Am., vol. 95, no. 5, pp. 2967, 1994

[55] Cook, P.R., "Physically inspired sonic modeling (PhISM): Synthesis of percussive sounds", Comput. Music J., vol. 21, pp. 38-49, 1997

[56] Rath, M. "An expressive real-time sound model of rolling". in *The 6th International Conference on Digital Audio Effects (DAFx-03)*. Queen Mary, University of London, 2003, 165-168.

[57] Stoelinga, C. and A. Chaigne, "Time-Domain Modeling and Simulation of Rolling Objects", Acta Acust United Ac, vol. 93, pp. 290 – 304, 2007

[58] van den Doel, K. "Physically-Based Models For Liquid Sounds". in *Tenth Meeting of the International Conference on Auditory Display*. Sydney, Australia, 2004.

[59] Avanzini, F., S. Serafin, and D. Rocchesso. "Modeling Interactions Between Rubbed Dry Surfaces Using an Elasto-Plastic Friction Model". in *Proceedings of the COST-G6 Conf. Digital Audio Effects (DAFx-02)*. Hamburg, 2002.

[60] Lakatos, S., P.C. Cook, and G.P. Scavone, "Selective attention to the parameters of a physically informed sonic model", J Acoust Soc Am, vol. 107, no. 5,Pt1, pp. L31-L36, 2000

[61] Lutfi, R.A. and C.-J. Liu, "Individual differences in source identification from synthesized impact sounds", J Acoust Soc Am, vol. 122, no. 2, pp. 1017-1028, 2007

[62] Reed, R.K. and G.R. Kidd, *Detection of Spectral Changes in Everyday Sounds.* 2007, Indiana University, Unpublished data.

[63] Fernström, M. and E. Brazil. "Human-Computer Interaction Design based on Interactive Sonification - Hearing Actions or Instruments/Agents". in *Proceedings of the 2004 International Workshop on Interactive Sonification*. Bielefeld University, Germany, 2004.

[64] Heller, L.M.W., L., "When sound effects are better than the real thing", J Acoust Soc Am, vol. 111, no. 5 pt.2, pp. 2339, 2002

[65] Moore, B.C.J., *Cochlear hearing loss*. Whurr Publishers, London, 1998.

[66] Badran, S. and E.L. Osama, " Speech and environmental sound perception difficulties by patients with hearing loss requiring and using hearing aid", Indian J Oto, vol. 4, no. 1, pp. 13-16, 1998

[67] Shannon, R.V., F.-G. Zeng, V. Kamath, J. Wygonski, and et al., "Speech recognition with primarily temporal cues", Science, vol. 270, no. 5234, pp. 303-304, 1995

[68] Reed, C.M. and L.A. Delhorne, "Reception of Environmental Sounds Through Cochlear Implants", Ear Hear, vol. 26, no. 1, pp. 48-61, 2005

[69] Inverso, D., *Cochlear Implant-Mediated Perception of Nonlinguistic Sounds*.Ph.D. Thesis Gallaudet University, April 2008

[70] Bronkhorst, A.W. and R. Plomp, "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing", J. Acoust. Soc. Am., vol. 92, no. 6, pp. 3132, 1992

[71] Festen, J.M. and R. Plomp, "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing", J. Acoust. Soc. Am., vol. 88, no. 4, pp. 1725, 1990

[72] Loizou, P.C., Y. Hu, R. Litovsky, G. Yu, R. Peters, J. Lake, and P. Roland, "Speech recognition by bilateral cochlear implant users in a cocktail-party setting", J. Acoust. Soc. Am., vol. 125, no. 1, pp. 372, 2009

[73] Cherry, C., "Some experiments on the recognition of speech with one and with two ears", J Acoust Soc Am, vol. 26, pp. 975-979, 1953

[74] Salthouse, T.A., "The processing-speed theory of adult age differences in cognition", Psychol Rev, vol. 103, no. 3, pp. 403-428, 1996

[75] Divenyi, P.L., P.B. Stark, and K. Haupt, "Decline of Speech Understanding and Auditory Thresholds in the Elderly", J Acoust Soc Am, vol. 118, pp. 1089-1100, 2005

[76] Humes, L.E., J.H. Lee, and M.P. Coughlin, "Auditory measures of selective and divided attention in young and older adults using single-talker competition", J. Acoust. Soc. Am., vol. 120, no. 5, pp. 2926, 2006

[77] Roth, P., L. Petrucci, T. Pun, and A. Assimacopoulos. "Auditory browser for blind and visually impaired users" in *CHI '99 extended abstracts on Human factors in computing systems*. Pittsburgh, Pennsylvania: 1999 of Conference.

[78] Mynatt, E.D., "Transforming graphical interfaces into auditory interfaces for blind users", Hum.-Comput. Interact., vol. 12, no. 1, pp. 7-45, 1997

[79] Cobb, N.J., D.M. Lawrence, and N.D. Nelson, "Report on blind subjects' tactile and auditory recognition for environmental stimuli", Percept Mot Skills, vol. 48, no. 2, pp. 363-6, 1979

[80] Petrie, H. and S. Morley. "The use of non-speech sounds in non-visual interfaces to the MS Windows GUI for blind computer users" in *Proceedings of the 5th International Conference on Auditory Display (ICAD98)*. University of Glasgow, U.K.: 1998 of Conference.