Evaluating non-speech sound visualizations for the deaf

TARA MATTHEWS*†, JANETTE FONG‡, F. WAI-LING HO-CHING§ and JENNIFER MANKOFF‡

†Computer Science Division, University of California, Berkeley, USA

‡Human-Computer Interaction Institute, Carnegie Mellon, USA

§Mckesson Medical Imaging Group, Richmond, CANADA

*Corresponding author. Email: tmatthew@cs.berkeley.edu

Sounds such as co-workers chatting nearby or a dripping faucet help us maintain awareness of and respond to our surroundings. Without a tool that communicates ambient sounds in a non-auditory manner, maintaining this awareness is difficult for people who are deaf. We present an iterative investigation of peripheral, visual displays of ambient sounds. Our major contributions are: (1) a rich understanding of what ambient sounds are useful to people who are deaf, (2) a set of visual and functional requirements for a peripheral sound display, based on feedback from people who are deaf, (3) lab-based evaluations investigating the characteristics of four prototypes, and (4) a set of design guidelines for successful ambient audio displays, based on a comparison of four implemented prototypes and user feedback. Our work provides valuable information about the sound awareness needs of the deaf and can help to inform further design of such applications.

Keywords: Sound visualization, peripheral displays, deaf

2000 Mathematics Subject Classification: 68N99; 94A99

1 Introduction

Sounds occur around us constantly, keeping us aware of our surroundings. Ambient sounds inform people of serendipitous events (co-workers socializing in the hall, music playing in a public place, children playing in the next room), problematic things (faucet dripping, fire alarm low-battery indicator, cell phone ringing at inappropriate times), and critical information (fire alarm, knocking on the door) relevant to their current situation or location. Maintaining this awareness is difficult for people who are deaf, though some tools exist to alert people who are deaf of particular events, such as the phone, doorbell, or fire alarm. However, there is no tool that provides continuous awareness of all the sounds in an environment. People who are deaf rely on vision, sensing vibrations in the ground, and other techniques to maintain awareness, but not every sound creates a vibration or leaves a visual trace. Sounds like knocking on a door or people approaching from behind can go unnoticed.

In our research we extensively explored the design of peripheral, non-speech sound visualizations (see Figure 1, 5, and 6). The goal of our design work was to answer the following questions with people who are deaf:

- What sounds are important to people who are deaf?
- Where is sound awareness important (*e.g.* at home, work, or while mobile)?
- What display size is preferred (*e.g.* a PDA, PC monitor, or large wall screen)?

- What information about sounds is important (*e.g.* sound recognition, location, or characteristics like volume and pitch)?
- What visual design characteristics are preferred?
- What functional issues are important in a visualization of non-speech sounds?
- How do different types of sound information effect user distraction and ability to identify sounds?

[Insert Figure 1 here.]

This paper builds on two pieces of previous work. One presents a rich set of design preferences and requirements based on interviews with participants who were deaf, culminating in a qualitative study of two displays exploring functional requirements such as sound recognition and history (Matthews *et al.* 2005). The other compares two other displays quantitatively along the dimensions of distraction and awareness, looking at the importance of different sound characteristics (Ho-Ching *et al.* 2003). This paper provides additional details about these past studies, and builds on them by presenting a set of design guidelines derived from both projects, shedding light on how to create a more informative and reliable peripheral display of sound that incorporates both computer and end-user interpretation of sounds.

After reviewing background work, we present a series of surveys and interviews in Section 3. First, we explored the sounds of interest to people in work, home, and mobile environments, ranging from serendipitous sounds such as children laughing to critical sounds such as fire alarms. Based on this, we generated design sketches that we presented in our interviews, leading to visual design preferences and functional requirements for peripheral visualizations of non-speech audio (one major contribution of this paper). Section 4 describes four prototypes that we implemented to explore different issues in studies: two address functional requirements (*e.g.* sound identification and history of past sounds) and the other two explore design dimensions related to distraction and awareness. Section 5 and Section 6 describe a second contribution of the paper: two lab studies of the prototypes. For one, we evaluated two fully functioning prototypes that we designed to embody the preferences and requirements gathered in interviews. In another lab-based evaluation, we explored how information about sound characteristics such as volume, pitch, and location affected a user's distraction and ability to identify sounds. In section 7, we present a set of design guidelines based on the preceding work. This represents the third contribution of our paper. Finally, we compare the capabilities, strengths, and weaknesses of each interface and conclude with a discussion of future work.

2 Background

Assistive technology for the deaf has focused on support for verbal communication. Common assistive technologies that focus on communication include assistive listening devices (these improve the audibility of one specific sound source that is likely to be lost due to distance or background sounds such as a lecturer in an auditorium or a conversation in a loud place), telecommunication devices (such as text telephones (TTYs) and video relay services), and close-captioning for television and movies (Mann and Lane 1995, Cook and Hussey 2002). In educational settings, classroom dialogue is captioned with educational transcription services, computer-assisted note-taking, captioning services, and, more recently, automatic speech recognition programs (Doyle and Dye 2002). There is a body of work focussed on automatic sign language recognition. Ongoing projects are developing techniques for capturing, segmenting (delimiting), and classifying sign language gestures. A summary of this work can be found in (Edwards 1997). Similarly, there are systems that perform machine translation of English into American Sign Language. A summary of work in this field can be found in (Huenerfauth 2003). Additionally, a number of systems have been developed to enable people who are deaf to practise articulation in speech therapy. One example is the auditory visual articulation speech

therapy software offered by Sonido (2006) that provides spectrograph visualizations of speech and allows drills against recorded speech samples. Work by Ellsman and Maki investigates the effectiveness of speech training using spectrographs compared to the effects of non-instructional training (Elssmann and Maki 1987). Their work suggests that spectrographs can enable students to drill on their own to a limited degree, and indicates that spectrographs might be one promising way to visualize non-speech audio. However they also discovered that certain speech sounds (/k/ vs. /t/) are not differentiated well on a spectrograph. One mobile system developed by Yeung *et al.* translated the fundamental frequency of speech intonation into vibrations in an array of solenoids worn on the wrist (Yeung *et al.* 1988).

Thus, past research looks at *spoken communication*, and past visualizations are intended for use in *focal* contexts. In contrast, our work focuses on *peripheral* awareness of *non-speech* sounds. Unlike verbal communication, sound awareness has received little attention in the research community to date. However, there are a variety of sound awareness techniques and products currently in use by people who are deaf. We present a review of these techniques, gathered from literature on assistance for the deaf (Mann and Lane 1995, Cook and Hussey 2002) and interviews with ten deaf participants, ASL interpreters, and an assistive technology consultant. A summary of our observations appears in Table 1.

In comparing these techniques, it is important to consider several dimensions of a technique:

- Is it interrupt-based, polling-based, or both? For example, visual inspection requires the user to regularly check on something (such as the toaster to see if the toast has appeared). This is a *polling* activity. In contrast, a hearing dog will *interrupt* the user, for example by nudging her when the phone rings.
- Does it support identification of ambient noise, notifications, or both? For example, hearing dogs will not always provide information about background noise (*ambient* sounds), while vibration sensing can provide information about ongoing ambient sounds (such as loud music). In contrast, a technique such as an alerting system is specialized to *notify* a user when something important happens (such as the phone or doorbell ringing).
- How high are the set-up and maintenance costs of the technique? Set-up and maintenance costs can have a big impact on adoption and continued use of techniques. An example of a technique with high initial investment is the flashing light alerting system for phones. Every phone in a house must be connected separately and the light must be visible in different rooms. An example of a technique with high ongoing maintenance is a hearing dog, which requires care throughout its lifetime.
- Must the technique be configured to a known, fixed set of sounds, or can it support awareness of unexpected or new sounds? For example, alerting systems are typically created to respond to a specific electronic or sound event, while vibration sensing through the floor can alert a user to unexpected or new sounds.

[Insert Table 1 here.]

In spite of the availability of these techniques, there is still room for improvement. Most of these techniques do not capture every sound in the environment. Vibration sensing is only effective for sounds that can be carried through the floor or through other physical media. Similarly, visual inspection is only effective for sounds that carry a visual component. Alerting systems only notify of events that have been hooked up previously and hearing dogs will not notify owners of every sound of interest. Although hearing aids do enhance all sounds, they are a high cost solution that requires training and they are not always effective.

In our present work we build on past work and available techniques, taking a user-centred approach to designing displays of sound for the deaf. We seek to answer questions about many design issues ranging from

place of use to type of information displayed. For example, is sound location more useful than sound volume and pitch? To gather this, we conducted an extensive design interview process, which resulted in new visual design knowledge and functional requirements that were important to the deaf participants. In the next section, we discuss our design interview process and results.

3 Gathering design requirements: survey and interviews

To inform the design of our applications, we conducted a survey and two sets of interviews. The initial survey and interviews enabled us to gather enough information to formulate concrete design questions. The second interviews with new participants explored the answers to these design questions. Specifically, we asked about

- place of use (home, work, mobile),
- size (PDA, PC monitor, large wall screen),
- type of sound information conveyed (sound recognition, location, and characteristics),
- visual design characteristics, and
- functional desires or issues.

3.1 Understanding awareness of sound in different settings

We began by surveying and interviewing ten hearing adults and ten deaf adults about the sounds they found most useful in various places and the techniques they used to stay aware of sound. Our goal was to generate a small set of diverse scenarios of different situations, not a comprehensive list of sounds of interest.

3.1.1 Hearing participants awareness of sound

We surveyed hearing people because they could be aware of sounds that the deaf might not notice. Thus, we gathered data from a small set of non-homogeneous hearing participants. Our group included a waitress, parking attendant, customer service representative, and a genetic researcher among other occupations.

Our participants were distributed a paper survey asking what sounds were important to them in their daily lives at home (they responded with television, phone, doorbell, alarms, talking, pets, cars, showers, music, e-mail / IM, wind, tapping, broken glass, alarm clock, answering machine, footsteps, microwave, cars, and construction) and work (they responded with boss, customers, co-workers, doors, phone / pager, cars, alarms, wheelchair motor, footsteps, printers, e-mail, and typing). Because our sample size was small, we cannot draw conclusions from the number of participants who mentioned each sound, but the list helps us better understand the design space. While analyzing the sounds participants found useful, we identified two types: *ambient sounds* and *notification sounds*.

Ambient sounds provided the participants with a general sense of what was happening in a space. They would continue in the background of their awareness, and would only come into the foreground when the participant wilfully concentrated their attention on the sound. Examples of these sounds include the television, showers, and music.

In contrast, *notification sounds* signified a particular event and required attention or action. These sounds gained a person's attention and distracted them from their current task. Examples of notification sounds include a telephone ring, doorbell, or alarm.

3.1.2 Deaf participants awareness of sound

To gather initial data on the tools and techniques in use by the deaf, we interviewed ten participants who were deaf and one assistive technology consultant. These interviews helped to inform the taxonomy of techniques presented in Table 1 (in our background section). Our interviews also helped us to learn about sounds

participants felt were inadequately supported by current techniques. In particular, users wanted an awareness of the following:

- *The activity and presence of others.* Having a sense of when they were alone was important for the participants we interviewed. Participants mentioned that they wanted to be aware of sounds like other people listening to music, expressly for this purpose.
- *Sound cues from appliances.* Participants described many instances of sound cues that they would like to be aware of. This is particularly important because many appliances such as kettles, microwave ovens, and smoke alarms are explicitly designed to use sound notification to keep users informed of important state changes, and often do not include adequate visual feedback. Others, such as faucets that drip when left on and printers that stop making noise when printing is complete have implicit sound cues that communicate their state.

Additionally, participants mentioned two situations in which current techniques were not adequate:

- *Away from the home environment.* A deaf person has more control over their home environment than over public environments such as the workplace. For example, they can choose a home with wooden floors which convey sound vibrations or invest in a light system to hook up to the doorbell. However, these systems are not commonly in place when the deaf person works amidst hearing co-workers or is in public areas. A few participants worked in an environment where the majority of the office workers were deaf. Their office space was designed so that everyone could see anyone at the doors and appliances were fitted with lights. However, such participants were the minority among those we interviewed.
- *In highly dynamic environments.* In an environment where needs change frequently, it is hard to correctly configure existing devices. For example, in an office where the hearing and non-hearing are mixed, hearing co-workers will often assist with certain sound awareness tasks such as receiving phone calls and visitors. However maintaining awareness of sounds when alone was a difficult problem for the deaf. One participant mentioned how waiting for a visitor can be very inconvenient. He would have to visually check every few minutes because he could not hear a door knock.

3.2 Needs and design sketch feedback

After gathering the data from our initial survey and interviews, we had a better understanding of our users' needs. With this understanding, we were able to define a set of questions for our second set of interviews that would help us develop scenarios and learn about users' design preferences for sound awareness tools:

- *Where is sound awareness important (if at all)?* We wanted to know more about situations that required increased sound awareness. For example, interviewees mentioned sounds in the home they wished to know about, but also expressed the need for sound awareness at work and in public places.
- *What size would users prefer (PDA, PC monitor, or large wall screen)?* Home, work, and public environments are very different in structure and use, calling for potentially different sized interfaces.
- *What type of sound information would be most useful?* Is it most useful for the system to show recognized sounds, where a sound occurred, characteristics of the sound (e.g. volume or pitch), or some combination of these three?
- *What visual characteristics (e.g. colours, shapes, icons, etc.) would enable people to best interpret sounds?*

- *What functional needs did users have?* For example, would users want a way to view the history of sounds that had occurred? What other functions should we incorporate into our designs?

To explore these questions, we conducted interviews that were split into two parts, occurring back-to-back. The first part was a formal interview. In the second part, we presented the participant with design sketches of potential applications and asked for feedback.

[Insert figure 2 here.]

3.2.1 Participants

We interviewed 8 participants who considered themselves deaf: two were profoundly deaf, two were mostly deaf, and four were hard-of-hearing with the help of hearing aids and mostly deaf without. Four participants wore a hearing aid(s), one had a cochlear implant, and three had neither. Three of our participants considered themselves culturally Deaf (see Senghas and Monaghan (2002) for a summary of what is known about deaf culture). We interviewed six females and two males, between the ages of 28 and 57. We had four participants who were employed full-time in an office, one student, one homemaker, one retired, and one unemployed.

3.2.2 Formal interview results

In the formal interview, we asked participants demographic questions, about the places where they spent time (e.g. home, work and other locations), the sounds of which they wanted awareness, and the tools and techniques they currently used to maintain awareness of these sounds. Most participants were interested in increasing their general level of awareness and were particularly concerned with alarms and other safety-related sounds. One participant was also very excited to learn about sounds. Though participants spent the majority of time at home or work/school, they emphasized a desire to be more aware of sounds in *all* places. In particular, they wanted to monitor sounds at work, home, in the car, and while walking.

At the office, participants wanted to know about the presence and activities of co-workers, emergency alarms, phone ringing, co-workers trying to get their attention, and faxes.

In their homes people were most interested in knowing about emergency alarms, wake-up alarms, doorbell and knocking, phone ringing, people shouting, intruders, children knocking things over, and appliances (faucets dripping, water boiling, the garbage disposal, gas hissing, *etc.*). One participant told a story about a time when his wife burned food, which caused the fire alarm to beep. Since both were deaf, they did not know the fire alarm was beeping until a hearing friend visited. Another participant expressed a desire to hear her children playing (or fighting) when she could not see them. A third participant told us, 'Once I left the vacuum cleaner on all night'. The same participant also told us that she needs a wake-up alarm because, 'Before an early flight, I will stay up all night'.

While mobile, participants were largely concerned about safety. While walking or running outside, people wanted to know about dogs barking, honking, vehicles, bikes or people coming up behind them, and whether they were blocking another person (e.g. 'excuse me', 'watch out'). One participant told about problems while running, 'When I first moved to L.A. I was surprised at how some drivers are aggressive on the roads and at intersections. I had some close calls'. While driving, people were interested in knowing about other cars honking, sirens, and sounds indicating problems with their vehicle. One participant told a story about how his car had developed a problem that, had he been able to hear the motor, he could have fixed before it caused serious damage, 'When there is something wrong with the car... it tends to go unnoticed until it is very expensive to fix'. Another participant said she used to own a convertible and would drive with the top down to be more visually aware.

When asked about current tools and techniques used for maintaining sound awareness, the results were similar to those found in our first set of interviews (presented in Table 1). In particular, all participants emphasized visual awareness: 'I tend to look forward ahead of me much further than typical people... My eye

sight is so important I've come to depend on it'. One new finding about tools and techniques in this second set of interviews was that most (5) participants did not have commercial alerting systems for the deaf (*e.g.* strobe lights for phone rings, doorbells, and emergency alarms). Participants explained that these tools were too expensive and difficult to install. Cost was also a major concern for all participants.

3.2.3 Design sketch interview results

Immediately following the formal interview with each participant, we began the design sketch interview. We presented participants with ten design sketches which were developed based on results from our formative survey and interviews. The design sketches, described in Table 2, represent variations on design characteristics relevant to our questions (place of use, size, type of information conveyed, visual design elements). We instructed participants to imagine that technical logistics (such as having an extra screen, battery life, and memory) are not a concern. For in-person and video relay phone interviews, each sketch and its description were on a single paper. For IM interviews, each sketch was shown on a webpage (<http://www.eecs.berkeley.edu/~tmatthew/ic2hear/>).

First, we asked participants to tell us their preferred place of use (*e.g.* home, office, or while mobile) and size (*e.g.* displayed on a PDA or PC screen) for each design. Second, we asked for feedback on the information the designs conveyed, which included recognized sounds, location of sounds, and sound volume and pitch. For example, Spectrograph with Icon (Figure 1b) showed recognized sounds (icons), volume, and pitch (spectrograph), while Rings (Figure 3a) showed location (position of the rings on the screen), volume (ring size), and pitch (ring colour). Finally, we asked about different visual design characteristics. For example, Rings (Figure 3a) used circular rings of varying sizes to indicate a sound's volume, and the colour of the rings to indicate pitch. Alternatively, Directional Icons (Figure 3b) used small iconic images to represent a sound. We also asked general questions about preference for and problems with designs and encouraged brainstorming about how designs could be improved. Table 2 summarizes our results.

Overall, participants tended to prefer the three displays that showed recognized sounds. Spectrograph with Icon (Figure 1a, for the implemented version) displays an icon or text of recognized sounds on a spectrograph (*e.g.* a phone icon appears over the spectrograph when a phone rings). Directional Icons (Figure 3b), shows icons at the edges of the computer screen, indicating both what sound occurred and its relative location to the screen. Map (Figure 2a), the other top choice, did not have recognition built in, but participants 'improved' it by adding recognition before selecting it. Map gives an overview map of a room and displays icons for recognized sounds or coloured rings for unrecognized sounds on the map where sounds occur.

Participants liked that these displays identified sounds of importance and conveyed this information with easy to understand icons. They felt that these three displays best allowed them to 'look to instantly know' or 'glance at it and figure out the sound'. Participants liked the location information in Map because it gave them even more information with which to identify sounds. Because of screen real estate concerns, participants liked the minimal, highly informative use of screen space of Spectrograph with Icon and Directional Icons. Two participants brainstormed a new display that showed a single icon (for recognized sounds) and rings (for unrecognized sounds) in the corner of a PC screen. The participants thought this would be a less distracting, smaller alternative to the other displays.

Participants did not like displays that they felt were harder to interpret (like Ambient Visualization: 'I'd have to practise and learn this to understand it'), less glanceable (like You Map: 'the Map is better... more clear'). Displays that showed location (making sounds easier to identify) tended to rate higher than those that did not. Participants also disliked displays that were inappropriately distracting (like Rings: 'Do you think someone is going to want to see those rings ALL the time on the monitor – that'd be annoying').

[Insert Table 2 here.]

Regarding size, participants preferred smaller displays in all locations, either on a PDA or using part of a PC screen. However, at home participants also valued large, wall screens for better visibility throughout a room. One participant described her ideal display at home, which was Map on a ‘flat panel LCD display that I can just stick to the wall in a visible location’.

Regarding display functionality, participants raised several issues. First, participants wanted a way to look at a history of identified sounds. One participant commented ‘I wouldn’t want to be looking at the monitor all the time, [but it would work] if it has a history component’. Another participant mentioned how a log could help: ‘[I] could see that [I] missed the phone ring’. Second, participants wanted the ability to customize which sounds were shown in order to manage distractions. For example, one participant wanted to change which recognized sounds were displayed (‘turn down the sensitivity of the display ...’) depending on context such as workload and amount of noise. Another wanted to minimize unimportant background noises (e.g. ‘I don’t really care about hearing the other environmental noises’). Four participants wanted to select which recognized sounds would be displayed. Third, two participants were sceptical about the accuracy of a computer system, showing concern about it displaying false information. One participant said ‘I am worried about it showing “voices” when I’m at home. If no one was there I would wonder “what is going on?”’ Another participant said, ‘I wouldn’t want to have it show a phone and then I look at my phone and it didn’t ring... I trust my own ability to interpret sounds more than the computer’s’. Clearly, the system’s confidence in the accuracy of displayed information needed to be conveyed to users.

[Insert Figure 3 here.]

3.2.4 Summary of interview results

The interview results provided us with an understanding of participants’ visual design preferences and functional requirements. Visually, participants preferred designs that were easy to interpret and glanceable. Displays with icons were preferred because participants could easily understand what sound occurred in a glance. Participants criticized displays they thought would be overly distracting, like Rings. More complex displays like Ambient Visualization were criticized for being difficult to understand. In addition, given the limitations of existing sound recognition technology, our results indicated the importance of enabling users to interpret sounds that are not recognized by a computer using features such as location. Participants tended to prefer displays that showed location or identity of sound over volume and pitch alone, although participants thought all features would be useful in identifying unknown sounds. Functionally, we found that participants wanted mechanisms to

- *identify* what sound occurred, *with or without* computer recognition,
- view a *history* of displayed sounds,
- *customize* the information that is shown, and
- determine the *accuracy* of displayed information.

[Insert Figure 4 here.]

4 Interfaces and implementation

In this section, we describe the four interfaces resulting from our design requirements. These include a display that arose due to participant brainstorming (Single Icon,* Figure 1b) one other popular display (Spectrograph with Icon, Figure 1a), and two displays that do not require a sound recognition system (Spectrograph, Figure 5, and Map, Figure 8). Taken together, these displays let us explore a range of issues that came up in our

* The Single Icon display was invented by two participants, though it was not one of our design sketches.

interviews including a variety of approaches to sound identification and a range of sizes. Additionally, all four displays were designed to help us get feedback through two lab studies described in sections 5 and 6. The first two displays were designed to help us explore semi-realistic use of displays with a full range of functionality (including sound recognition, history, customization and accuracy) in a qualitative study. The second two displays were selected to give us more information about design dimensions that are difficult to study through interviews, such as distraction from a primary task, ability of different visual features to support correct detection of target sounds, and performance in noisy and quiet environments. They were compared in a controlled study. Implementation decisions described below reflect these different intended uses.

4.1 Single Icon

The Single Icon prototype (Figure 1b) is a minimalist display proposed by participants in our interviews that supports all four functional requirements (sound identification, confidence, history, and opacity). It graphically displays recognized sounds as icons and unrecognized sounds as rings. It has a small footprint (55 x 93 pixels) that conserves screen space and lessens distraction. Ring colour represents pitch (red for high, blue for low) and the number of rings represents its volume (many rings for loud, few rings for soft). Icon opacity indicates confidence, along with the words 'High' 'Medium' or 'Low'. Users can also select which sounds to display. The prototype can optionally show a history of past recognized sounds, each represented as a coloured bar in a vertical bar graph (Figure 4). On this History Display, the x-axis represents time and the y-axis represents the volume of the sound.

Single Icon was implemented using the Peripheral Display Toolkit (Matthews *et al.* 2004), a toolkit implemented in Java to support the creation of peripheral displays. Sound recognition was handled by Malkin's state-of-the-art system (Malkin *et al.* 2005), which uses audio only to detect and classify events. We chose (Malkin *et al.* 2005), an audio-only system, because of the prohibitive cost, complexity, and intrusiveness of the many sensors involved in multimodal event detection (see Oliver and Horvitz (2002), for example).

The recognition system can be trained for use in any place. To capture audio, we used the Sony ECM 719, a high-quality, one-point stereo recording microphone. The sound recognition system returns a best guess as to the classification of a sound, with a normalized confidence level based on thresholds that are set based on test data.

4.2 Spectrograph with Icon

As with the previous prototype, Spectrograph with Icon fulfils all four functional requirements. It includes the same History Display and confidence display as the previous prototype, which it expands on. Spectrograph with Icon adds a black and white spectrograph to the previous prototype. This increases its footprint to a conservative 263 x 155 pixels. Spectrograph with Icon shows more detailed information about sounds than the previous display. It provides high-fidelity information about volume (brightness) and pitch (vertical axis) over time (horizontal axis), helping users to interpret sounds. The spectrograph enables users to participate in sound identification and be more aware of unrecognized sounds, and can scaffold a user in learning about or discovering sounds. For example, mechanical sounds are easy to visually identify as they often have regular pitch and amplitude patterns. Also, spectrographs have been used in the past for speech therapy, where sound identification is important (Elssmann and Maki 1987).

Implementation was identical to Single Icon, with the addition of a Spectrograph window. The spectrograph window was a build-in part of Malkin's sound recognition system, so the same audio input that went to the recognition system also went to the spectrograph. We modified it to occupy less screen space and to be positioned next to the icon window.

[Insert Figure 5 here.]

4.3 Spectrograph (without icon)

The Spectrograph prototype (Figure 5) is visually similar to Spectrograph with Icon, except with the addition of colour (which enhances the information about volume). Because this prototype is intended for use in a controlled, quantitative lab study, history and customization were not supported. Instead, this prototype was designed to allow us to study certain visual design characteristics by comparing it with the next prototype, Map. We implemented this prototype using a scripting language, Python, with the SNACK v2.2a1 toolkit (Sjölander 2002), which provides facilities for manipulating real-time sound data, including a configurable spectrograph window, which we used to implement this prototype. This combination let us easily control exactly what audio data was sent to the prototype, enabling us to repeat identical visual stimuli with each participant in our lab study. Because this display is not highly glanceable, we expected it to be more distracting than the Map display.

[Insert figure 6 here.]

4.4 Map

In the Map prototype (Figure 6), the background displays an overhead map of a room and sounds are depicted as rings. The centre of the rings denotes the position of the sound source in the room. The size of the rings represents the amplitude of the loudest pitch at a particular point in time. Each ring persists for three seconds before disappearing. Unlike our design sketch, this prototype does not incorporate frequency information. Map allows for limited identification of unknown sounds based on position and amplitude. Map was intended as a glanceable secondary display, though the lab study would investigate how well and with how much distraction it conveyed information. Again, it was unnecessary to incorporate support for history (though it does show three seconds of data) or user customization given the goals of our lab study. As with the Spectrograph, this prototype was designed for a controlled study, and it was implemented using the same tools (Python and SNACK). For this reason, the map was hard-coded to match the space in which we would conduct the study. Also, our study used Wizard-of-Oz to enter sound location (a full implementation would require additional software and hardware, including either a microphone array (Asano *et al.* 2000, Bian *et al.* 2005) or a PC with multiple sound cards (Scott and Dragovic 2005)). In particular, we provided the Wizard with a second window showing the map. Each time a sound occurred, the wizard clicked in the appropriate location on his or her map. The user's map would then show the sound (with dynamically calculated ring sizes) in the location indicated by the Wizard.

4.5 Summary of four displays implemented

In summary, we implemented four different displays, intended to contribute to two different evaluations that would complement our design interviews. Table 3 compares the characteristics of the four displays: what type of information they display (position, amplitude, frequency, recognized sounds), whether the prototype requires training (*i.e.* training the system to recognize sounds), requires Wizard of Oz input, enables user customization, indicates its confidence in displayed information, and shows a history of sounds.

[Insert table 3 here.]

We conducted two independent studies. One was a qualitative lab study comparing the two sound recognition interfaces in semi-realistic use, meant to gather user feedback on how well the designs conveyed the right information in an office setting (for recognized and unrecognized sounds), and satisfied functional requirements (*i.e.* confidence, customization, history). Because the use of these prototypes was open-ended, we made sure that they supported all of the functional requirements (see the first two interfaces listed in Table 3). A second, controlled quantitative lab study comparing the interfaces without sound recognition, focused on answering two additional questions: (1) can the user learn to interpret sounds without the computer recognizing them; and (2) would the peripheral displays be too distracting, or require too much of the user's attention. To help answer these question in a study, notice that the third and fourth prototypes in Table 3 differ in the information they convey and do not identify sounds for users. Next we present both studies (Sections 5 and 6, respectively), followed by Section 7, a discussion section comparing and contrasting what we learned across both studies.

5 Evaluating the sound recognition interfaces

To evaluate our implemented sound recognition applications, we asked four people who were deaf to use each of the two sound recognition displays and give us feedback. Our goal in running this study was to get qualitative feedback allowing us to compare the two designs. In particular, we set out to answer the following questions: (1) how well do the designs convey the right information in an office setting (for recognized and unrecognized sounds); (2) how well do the displays satisfy functional requirements (*i.e.* confidence, customization, history); and (3) overall, do users consider the displays useful. Our results showed that the displays were considered useful, that history was a highly valued part of display functionality, and that sound recognition is helpful and preferred but would be difficult to train for all sounds of interest to users.

We wanted to observe the participants' use of the tool in as realistic an environment as possible, to help participants imagine how such a tool could be used. Again, our ultimate goal is to create a system that could be evaluated in the field, so realism was important. Since we were testing the system in our lab, we chose to focus on an office environment, simulating events that were of interest to participants in our earlier interviews: speech, phone ringing, door opening/closing and knocking. We also trained the system on common office sounds to be filtered out: typing, mouse clicks, chair creaks, and continuous background noises (*e.g.* heaters and fans).

5.1 Participants

Each participant had a varying degree of hearing ability. The first participant we interviewed was profoundly deaf. The second wore a hearing aid to hear sounds but could rarely identify them. An ASL interpreter served as a translator for these two participants. The third and forth participants wore hearing aids and could hear and identify many sounds with them. During the study, both participants 3 and 4 could hear most sounds and identify some, such as a phone ring. The participants included a technical support worker, a journalist, a substitute teacher/graduate student, and a childcare worker.

5.2 Results

Variation in hearing ability among participants affected their preferences, though all said that sound recognition was important to them. All four participants favoured the History Display as a stand-alone display. Overall, participants were positive about the applications because '[people who are deaf] miss out on a lot of things'.

5.2.1 Reactions to Single Icon

All participants liked Single Icon because it identified each sound event. One participant said, '[I] need to know what the sound is'. One participant was especially pleased when he saw the phone icon and turned to see the researcher still holding the phone, verifying the system's sound identification.

Participants told us which display aspects could be improved. Two participants wanted the icon to change less often, as its frequent changing had confused and distracted them at inappropriate times. Also, they requested visual cues as to which sounds were really important (*e.g.* flashing the icon). For less important sounds, they wanted more 'visual quiet'. Two participants emphasized the importance of knowing the location of a sound not recognized by the system, and one participant wanted to know the number of times it had occurred. The same participant emphasized that she already identifies sound 'using process of elimination', based on where the sound occurred. One participant who could hear some sounds was frustrated when the system reported a sound inaccurately, or when a sound was correctly identified but the system reported a 'Low' confidence. In general, participants disliked that the icon's opacity varied with the system's recognition confidence. A few participants complained that they could barely see the icon when it was less opaque.

Participants appreciated the ability to turn off notification of certain sounds, and wished for a higher level of customization. While playing with the interface, one participant turned off the phone ring notification because 'I'd see the phone and wonder whose it is... [I'd be] picking up all the phones [in my office]'. Participants also wanted to be able to change the size of the display. For example, one participant wanted the display to be smaller, while another wanted it to be larger. One participant wanted the ability to hide the display completely.

Three of the four participants were interested in having the Single Icon display implemented for their pagers. The participants emphasized that their PDAs were their most reliable connection to the outside world. Two participants emphasized their use of the instant messenger program on their PDAs and viewed them as an alternative to their computers. All participants emphasized that they would not be sitting in front of their computers all day and would need an alternative type of display, such as on their PDA or television. One participant said, 'I move around [and] am not always at my desk'. One of the participants mentioned that Single Icon on a PDA would be useful for people with kids, saying that '[I had a] baby light for the baby, but every time I moved to another room I had to set it up [again]'

5.2.2 Reactions to Spectrograph with Icon

In general, Spectrograph with Icon was liked less by the participants. Participants often did not understand what the spectrograph was actually showing them. One participant joked, 'There are ghosts on your computer'. All participants were doubtful that they could learn to recognize the patterns on the spectrograph. One participant said, 'It's just a blurry picture. I have no idea what the sound is'. Participants also felt that the spectrograph was too distracting. For example, a participant said, '[The spectrograph] is like a hearing aid. It's just a bunch of noise, and I can't focus on what's important'. One participant suggested that the icon and the spectrograph appear in the same window to minimize the distraction.

However, three out of four participants expressed interest in using Spectrograph with Icon to learn about sounds, playing with it by producing sounds and watching the spectrograph output. One participant watched the spectrograph to see if she could learn to identify a sound by its shape, saying 'I wanted to see what [the sound] was. I wanted to see the shape on the spectrograph to see if I could recognize it'. In addition, the spectrograph was effective at attracting participants' attention when a loud sound occurred. For example, when the door was slammed, a participant looked at the display because '[the spectrograph] caught my attention because of the black'.

Because the spectrograph was presented as a method of visualizing non-recognized sounds, it prompted participants to think of alternative methods of event detection other than sound. One participant told the researcher about the various sensors she had put around her house. Another participant suggested that the

tool record sounds so that she could send it to a hearing person to identify it for her. This type of feedback further emphasized that the display must either identify sounds or help users identify sounds for themselves.

5.2.3 History Display

The History Display received the most favourable feedback from our participants, many of whom suggested that it be a stand-alone display. One participant explained that the History Display was especially useful because he did not have to watch it constantly, but he was still aware of the sounds around him. The same participant also mentioned incidences where he had woken up during the night, or felt something and wondered what happened, emphasizing the need for a display of recent sounds. Users found it particularly useful that the display showed relative volumes of sounds (*e.g.* louder sounds are depicted with taller bars). One user said, ‘small sounds I don’t bother with, but loud sounds [get my attention]’. All participants wanted to keep the History Display open, but make it ‘small and thin and have it at the top of my screen’. One participant contrasted this display with Single Icon saying, ‘[The History Display] is better, because the different pictures are not helpful’.

All participants also expressed the desire for the History Display and Single Icon to identify a larger set of sounds. The childcare worker and journalist/mother both expressed the desire to know about the sounds their children were making in other rooms. Another participant wanted the displays to show every sound that was made, including co-workers coughing or sneezing. This feedback indicates that while there may be a small set of common sounds in which many people would be interested, each individual will have varying preferences about the types of sounds of which they wish to be aware. Ideally, the tool would be flexible enough to allow users to be aware of any type of sound.

5.3 Study limitations

The key results of our study were that the displays were considered useful, that history was a highly valued part of display functionality, and that sound recognition was helpful and preferred but would be difficult to train for all sounds of interest to users.

These results must be viewed within the limitations of this study. First, we had only four participants, which reduces the generalizability of the results. Second, Spectrograph with Icon was consistently presented after Single Icon, which may have biased users. Finally, participants used the displays for a short time in the lab (30 minutes each), which is unlikely to produce results that are highly comparable to long-term, real-world usage.

6 Evaluation of correct identification and distraction

In a lab study evaluating our two displays of unrecognized sounds, Spectrograph (without icons) and Map, we set out to answer the following questions about identifying sounds:

- Which display allowed for more correct identifications?
- How does the presence or absence of background noise affect the identification of sound?

We sought answers to the following questions about distraction:

- How distracting were the displays?
- Was one display more distracting than the other?
- Were the displays *perceived* as distracting?
- Was one display *perceived* as more distracting than the other?

Our results helped to verify that participants were able to detect and correctly identify sounds using the displays. Map afforded significantly more correct identifications during Noisy trials. During Quiet trials, Spectrograph enabled marginally more correct identifications (with significance). Our results also show that

neither display was measured as significantly distracting. However, we found that participants were significantly more distracted in Noisy conditions than in Quiet conditions. Additionally, participants were significantly more distracted when using Spectrograph in Noisy conditions, than when using Map in Noisy conditions. Qualitatively, users preferred Map, though they liked both displays.

Because we wanted to study user identification of sounds and distraction, our study design was significantly different from the previous study. We used dual-task methodology to study distraction. As a consequence, our study was primarily a quantitative study. We were also able to gather qualitative data during interviews after the study trials and during the training period.

6.1 Method

Each session was held in a quiet room resembling an office. Two monitors were connected to a single computer running the study software. The study had a 2 x 2 x 2 within subjects factorial design with a dual attention task. We manipulated three independent variables: prototype (Spectrograph, Map), background noise level (Quiet, Noisy) and notification sound (Door Knock, Phone Ring). The background noise was speech, produced by a random selection from five different speech files. The notification sounds were chosen to represent common notification sounds.

All sounds in this study were pre-recorded and we did not use a microphone to detect naturally produced sounds. Instead, sounds were recorded and played at random intervals by a computer program. Audio output from the program was fed directly into the input of displays. Thus every door knock produced exactly the same sound signature as every other door knock with a consistent frequency mix and amplitude. This was desirable as we were conducting a controlled study and wanted to exclude any extraneous variables from confounding our data. This included background noise naturally occurring in the room and any variation in frequency or amplitude in the target sounds.

We used a dual-task methodology to measure distraction with each task being presented on a separate monitor. The primary task was a visual task in which participants searched a screen of numbers in the range [0, 9]. Each participant was told to find and click on as many 0s as possible. After a timer expired, the trial would end and a new number field would be generated for the next trial.

The secondary task appeared on a second monitor and required the participant to identify one of two sounds, a door knock or a telephone ring. Each participant was told to press the <ESC> key on the keyboard when one of the notification sounds was identified. A dialog box would then appear that would ask participants to indicate which sound they had identified (Phone Ring or Door Knocking) and with what certainty on a 5 point scale (1=Unsure, 5=Very Sure).

Sounds were presented to the participants in a series of blocks of trials. During a block, the participant would perform the primary task, searching the number field on a primary monitor. At the same time, s/he would perform the secondary task, to monitor the secondary screen for a notification sound. A notification sound was produced at a random time. The appearance of the notification sound would end the trial. Immediately after the sound was played, the system would time out and the next trial would start after the time out period had expired, unless the sound was correctly identified. In this case the next trial would begin as soon as the participant had indicated that he had heard the sound.

Further details of the experimental method can be found in [6].

6.2 Participants

There were eight deaf participants, five male and three female. All were adult office workers. None had non-corrected visual impairments, such as colour blindness. All were profoundly deaf. One regularly wore a hearing aid, but removed it for the session. All had computer experience through their work. Instructions for

the study were available in PowerPoint slides, which the participants read on their own. Participants had the choice of communicating with the researchers through an ASL interpreter or with pencil and paper.

6.3 Results

Participants reacted favourably to both prototypes, though all eight participants preferred the Map visualization to the Spectrograph. During the pilot studies, they would experiment by clapping their fingers, jingling keys and making sounds with their voices to see the effect on the visualization. As one participant put it, ‘This is great! ... I’m learning to hear again after 30 years’!

Quantitatively, we compared the two visualizations (Map, Spectrograph) in different environments (Noisy, Quiet) in terms of correct identification of signal sounds, distraction from the primary task, and learning (using a multivariate analysis of variance, or MANOVA). All results presented are significant. Main effects were found for all three independent variables (display, environmental conditions, and sound played) and interaction effects were found for the display * environmental condition (Wilks’ $\Lambda = .908$, $F(5,376) = 7.6$, $p < .001$), and for all three independent variables (Wilks’ $\Lambda = .818$, $F(10,752) = 2.9$, $p < .001$).

[Insert figure 7 here.]

6.3.1 Identification

With regard to correct identifications, Map performed significantly better in Noisy conditions and Spectrograph performed marginally better in Quiet trials (with significance). Overall, however, the Map display performed better across the different conditions. These effects are illustrated in Figure 7.

Users of Spectrograph were able to detect sounds of interest an average of only 50% of the time, with a certainty of only 2.7 on a 5 point scale, in the Noisy condition. In contrast, participants had average identification rates of 80% or higher when using the Spectrograph in the Quiet condition, or the Map display in either condition. Users of Spectrograph were significantly more sure of their answers than users of Map in the Quiet condition ($M_{cert} = 4.4$ versus $M_{cert} = 3.6$), while users of Map were significantly more sure of their answers in the Noisy condition ($M_{cert} = 4.2$ versus $M_{cert} = 2.7$).

False alarms (thinking a sound had occurred when it had not) were not a large problem, with a median of 0 in both Noisy and Quiet conditions for both Map and Spectrograph. Qualitatively however, we observed that Spectrograph did not perform well for adjacent sounds. When a sound was played right after noise, participants told us they were unable to visually distinguish the sound from the background noise. However, a more nuanced understanding of false alarms shows that they were most likely to occur among Spectrograph users responding to knocks, followed by rings, and then no sound, in the Noisy condition ($p < .05$). This significant interaction effect between all three independent variables is illustrated in Figure 8. The Map visualization did not suffer from this problem since position disambiguated adjacent sounds.

[Insert figure 8 here.]

6.3.2 Distraction

Distraction was measured as the rate at which a participant could detect 0s in the primary task and the rate at which errors were made. We performed a one-way repeated measures ANOVA on primary task performance with and without a secondary display[†]. Neither secondary display was significantly distracting ($F_{obs} = 1.152$, $p > 0.05$).

[†] There were unequal sample sizes because two participants did not complete the baseline block, so we removed them from this analysis.

However, when we separated the trials into Noisy and Quiet conditions, we found that participants were significantly more distracted (*i.e.* they found fewer zeros per second) in Noisy conditions than in Quiet conditions using either display. Additionally, participants were significantly more distracted when using Spectrograph in Noisy conditions than when using Map in Noisy conditions. These effects are illustrated in Figure 9.

Analysis of the questionnaire suggested that Map was perceived as less distracting than Spectrograph. Although this difference was not significant ($p > 0.05$), visual inspection showed that seven of the eight participants reported the Map visualization as less distracting.

[Insert figure 9 here.]

6.3.3 Learning

Results from a post-study questionnaire reported that participants perceived Map as significantly easier to learn than Spectrograph. Each individual rated Map higher in ease of learning than Spectrograph with means of 5.43 and 3.57 respectively on a 7-point scale (1 = Difficult to learn, 7 = Easy to learn). This difference was significant ($p < 0.05$). There did not appear to be large learning effects during the trials. There were two blocks of trials per interface and the order of the blocks was not a significant factor towards the percentage of correct identification, according to a two-way repeated measures ANOVA ($F_{\text{obs}} = 1.795$, $p > 0.05$).

6.4 Study limitations

The key results were (1) the addition of the displays did not significantly distract from the primary task in general, though there was significant, measurable distraction in the noisy condition, especially among users of the Spectrograph; (2) participants detected a significantly higher percentage of sounds when using Map in Noisy conditions than when using Spectrograph in Noisy conditions; and (3) Spectrograph enabled a slightly more correct identifications in Quiet trials (with significance). In addition, participants perceived Map as easier to learn, although our quantitative data showed no significant difference between Map and Spectrograph in terms of learning.

These results must be viewed within the limitations of the study. First, though Map visualized position and amplitude and Spectrograph visualized frequency and amplitude, these displays represent two of many possible visualizations of those data dimensions. Different visualizations of the same data could very well have shown different results. Second, we only modeled a very small sample of sound identification situations that would appear in a real office. For example, our speech always originated from a single location, but speech in an office could originate from multiple locations, it could appear at the same location as the telephone or the door, or it could travel as two people talk while walking. Third, we only modeled two signal sounds (phone and knocking), but in reality there are many more interesting sounds to people who are deaf. All of these variables could produce different results than what was shown in this study. Fourth, the sounds in our trials occurred more often than in a typical office setting. We modeled an extremely busy office! Finally, though we did not measure any significant distraction, participants may have been in a heightened state of awareness as compared to a typical office worker.

7 Discussion

Here we combine results from design interviews and prototype evaluations to answer our research questions about the design of peripheral visualizations of non-speech sound for the deaf. In addition, we present two high-level guidelines for sound visualization applications that we discovered and confirmed over the course of

our research: (1) designs should identify or help the user identify sounds and (2) allow users to choose which sounds to show, filtering out the rest.

Is sound awareness important? If so, where is it important? Our participants confirmed that sound awareness was important, citing numerous examples in work, home, and mobile settings involving social interactions (e.g. presence of co-workers or children playing in another room), safety (e.g. fire alarms, intruders, and traffic when walking outside), and many other situations.

What display size is preferred? In all locations, participants preferred smaller-sized displays. However, in the home participants also valued large, wall screens for better visibility throughout a room.

What information about sounds is important to people who are deaf? Ideally, participants wanted to know exactly what sound had occurred, information that sound recognition can help provide. Almost all participants told us sound identification was critical to their ability to maintain awareness and react appropriately. For example, one user said that unidentified sounds when she is home alone are ‘scary’. If the system could not recognize the sound, users wanted information such as location or volume to help them identify it. Our combined evaluations explored the effectiveness of each of these sound information types, showing that each type of sound information is limited by itself, but a combination of sound recognition, location, volume, and frequency might improve sound identification in more situations. For example, without position information, participants often confused a door knock with background speech or the sound of their keyboard. Additionally, we found that a history of sounds was critical, because it allowed participants to maintain awareness of sounds without constantly watching the display.

What visual design characteristics are preferred? Participants preferred visual designs that were easy to interpret, glanceable, and appropriately distracting, over designs with more detailed information or one type of notification (e.g. only minimal distractions).

Results from our first prototype evaluation, clearly show that users preferred the History Display over Single Icon and Spectrograph with Icon. The reason is that it made better use of visual design characteristics: the History Display appropriately attracted user attention by showing relative volumes, unlike Single Icon which drew no distinction between more and less important sounds. Participants still liked Single Icon for its easy to interpret icons, which abstract sound identification and confidence levels into one picture. The less popular spectrograph was regarded as difficult to interpret and peripherally monitor because it displayed too much detail. Although previous work has documented that spectrographs have been used successfully as a focal display in speech therapy [6], different visual characteristics are desirable when using a spectrograph as a secondary display. Likewise, the second prototype evaluation showed that users preferred Map over Spectrograph because it was easier to interpret position than amplitude and frequency.

What functional issues are important? Important functional requirements identified in design interviews include the ability to identify what sound occurred, view a history of displayed sounds, customize the information that is shown, and determine the accuracy of displayed information. The evaluation of our displays validated that these issues were critical. For example, results from the first prototype evaluation showed that users appreciated the ability to turn off notification of certain recognized sounds. Results from the second evaluation showed that the users had a hard time identifying sounds with background noise present using Spectrograph. This display might have performed better if background noises had been filtered out.

How do different types of sound information effect user distraction and ability to identify sounds? While we cannot generalize to all displays of using position, amplitude, and frequency, we have a better understanding of how these types of information can effect identification and distraction. In Noisy situations, users correctly identified sounds better and were less distracted with position and amplitude (Map) than with amplitude and frequency (Spectrograph). Based on our observations, position was key to enabling sound identification in Noisy situations. In Quiet situations, amplitude and frequency (Spectrograph) were slightly better for identification than position and amplitude (Map).

In addition to these research questions, we have discovered and confirmed two high-level guidelines for an ambient sound visualization display:

1. *The display should identify or help the user identify sounds.* Almost all participants told us sound identification was critical to their ability to maintain awareness and react appropriately. For example, one user said that unidentified sounds when she is home alone are ‘scary’. Because sound identification was the major user goal, sound recognition received a great deal of positive feedback. However, due to limitations of sound recognition technology, recognition for all sounds is improbable. Users were willing to interpret sounds themselves, as long as the system could provide information that would help them do so. Our combined evaluations showed that each type of sound information is limited by itself, but a combination of sound recognition, location, volume, and frequency might improve sound identification in more situations.
2. *The display should allow users to choose which sounds to show and filter out the rest.* Ambient sounds are present all around us all the time. However, different sounds are of varying importance to a person depending on his/her context. Users need the ability to choose which sounds should be displayed. For example, results from the first prototype evaluation showed that users appreciated the ability to turn off notification of certain recognized sounds. Results from the second evaluation showed that the users had a hard time identifying sounds with background noise present using Spectrograph. This display might have performed better if background noises had been filtered out.

8 Conclusion and future work

In this paper, we presented an iterative investigation of peripheral, visual displays to help people who are deaf maintain an awareness of non-speech sounds in the environment. Our major contributions are: (1) a rich understanding of what ambient sounds are considered useful by people who are deaf; (2) a set of visual preferences (ease of interpretation, glanceability, and appropriate distractions) and functional requirements (the ability to identify sounds, view a history of displayed sounds, customize the information that is shown, and determine the accuracy of displayed information) for a peripheral sound display, based on feedback from people who are deaf; (3) lab-based evaluations investigating the functional characteristics of four prototypes with participants who are deaf; and (4) a set of design guidelines for a successful display of ambient audio, based on the comparison of four implemented prototypes and user feedback.

Our design interviews and evaluations have left us with a wealth of knowledge with which to design future applications. While the applications presented in this paper were deployed in an office setting, participants also expressed a need for sound awareness at home and in mobile settings. We plan to implement applications designed for home and mobile use. In addition, participants thought the location of a sound could be valuable for identifying it. We also plan to incorporate support for locating sounds in a space. Finally, we plan to conduct a long-term deployment of our iterated application.

9 References

1. Auditory Visual Articulation Therapy Software. Sonido Incorporated, Available online at: <http://www.sonidoinc.com> (accessed 12 January 2006).
2. ASANO, F., ASOH, H. and MATSUI, T., 2000, Sound source localization and separation in near field. *IEICE Transaction Fundamentals*, **E83-A** (11). pp. 2286-2294.
3. BIAN, X., ABOWD, G.D. and REHG, J.M., 2005, Using Sound Source Localization in a Home Environment. In *Proceedings of The 3rd International Conference on Pervasive Computing*, pp. 19-36.

4. COOK, A.M. and HUSSEY, S.M., 2002, *Assistive technologies: principles and practice*, (St. Louis: Mosby, Inc.).
5. DOYLE, M. and DYE, L., 2002, Mainstreaming the student who is deaf or hard-of-hearing. Alexander Graham Bell Association for the Deaf and Hard-of-Hearing, Available online at: <http://www.agbell.org/docs/mainstreamingDHH.pdf> (accessed 12 January 2006).
6. EDWARDS, A.D.N., 1997, Progress in Sign Language Recognition. In *Gesture and Sign-Language in Human-Computer Interaction*, (Springer), pp. 13-21.
7. ELSSMANN, S.F. and MAKI, J.E., 1987, Speech spectrographic display: use of visual feedback by hearing-impaired adults during independent articulation practice. *American Annals of the Deaf*, **132** (4). pp. 276–279.
8. HO-CHING, F.W.-L., MANKOFF, J. and LANDAY, J.A., 2003, Can you see what I hear? the design and evaluation of a peripheral sound display for the deaf. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI)*, (ACM Press), pp. 161-168.
9. HUENERFAUTH, M., 2003, Survey and critique of ASL natural language generation and machine translation systems. *Technical Report MS-CIS-03-32*, Computer and Information Science, University of Pennsylvania.
10. MALKIN, R., MACHO, D., TEMKO, A. and NADEU, C., 2005, First evaluation of acoustic event classification systems in CHIL project. In *Hands-Free Speech Communications and Microphone Arrays (HSCMA) Workshop*.
11. MANN, W.C. and LANE, J.P., 1995, *Assistive technology for persons with disabilities*, (Bethesda, MD: The American Occupational Therapy Association, Inc.).
12. MATTHEWS, T., DEY, A.K., MANKOFF, J., CARTER, S. and RATTENBURY, T., 2004, A toolkit for managing user attention in peripheral displays. In *Proceedings of the 17th annual ACM symposium on User interface software and technology (UIST)*, (ACM Press), pp. 247-256.
13. MATTHEWS, T., FONG, J. and MANKOFF, J., 2005, Visualizing non-speech sounds for the deaf. In *Proceedings of ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)*, pp. 52-59.
14. OLIVER, N. and HORVITZ, E., 2002, Layered representations for human activity recognition. In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces (ICMI)*, (IEEE Computer Society), 3.
15. SCOTT, J. and DRAGOVIC, B., 2005, Audio location: accurate low-cost location sensing. In *Proceedings of The 3rd International Conference on Pervasive Computing*, pp. 1-18.
16. SENGHAS, R.J. and MONAGHAN, L., 2002, Signs of their times: deaf communities and the culture of language. *Annual Review of Anthropology*, **31**. pp. 69-97.
17. SJÖLANDER, K. Snack toolkit, Available online at: <http://www.speech.kth.se/snack/> (accessed 12 January 2006).
18. YEUNG, E., BOOTHROYD, A. and REDMOND, C., 1988, A wearable multichannel tactile display of voice fundamental frequency. *Ear Hear*, **9** (6). pp. 342-350.

Figure 1

Table 1

Technique	Application	Characteristics	Cons	Interrupt	Polling	Ambient	Notification	Pre-configured sounds	Costs
Vibration sensing	Awareness of sounds that create vibration (<i>e.g.</i> sensing footsteps, feeling that a computer is on)	Does not require attention focus if individual is in direct contact with vibrating object	Depends on infrastructure (<i>e.g.</i> having hardwood floors)	Y	Y	Y	Y	N	L
Visual inspection	Multiple applications, <i>e.g.</i> kettle steam, looking out window for a guest's arrival	Sometimes is the only alternative	Different for each sound; Depends on polling	N	Y	Y	Y	N	L
Alerting devices / systems	Flashing lights, vibration, or extra-loud sounds provide awareness of notification sounds (<i>e.g.</i> phones, doorbells, babies crying, alarms, <i>etc.</i>)	Interrupt-based	Must install each device; Flashing lights require visual attention; User may have to wear a vibrator; Extra loud sounds affect others who are hearing	Y	N	N	Y	Y	M
Hearing dogs	Awareness of all sounds	Interrupt-based	Requires ongoing maintenance; Requires a priori training per sound.	Y	N	Y	Y	Y	H
Hearing aids, cochlear implants	Enhancing existing hearing, but not at the fidelity of "normal" hearing	Enhances awareness of all sounds	Requires training for interpretation of sounds; Results vary by case	Y	N	Y	Y	N	H

Figure 2

Table 2

Sound Info	Size		Name	Description
Recognized sound, volume, pitch	Small	👍	Spectrograph with Icon	An icon or text of recognized sounds is displayed on a spectrograph (<i>e.g.</i> , a phone icon appears over spectrograph when phone rings) (see Figure 1a for the implemented version).
Recognized sound, location, volume, pitch	Small	👍	Directional Icons	Icons appear at edges of computer screen, indicating both what sound occurred and its relative location to the screen (see Figure 3b).
	Large	👍	Map	Overview map of a room, icons (recognized sounds) or coloured rings (unrecognized sounds) appear on map where sound occurred (see Figure 2a).
Location, volume, pitch	Large	👉	Ambient Visualization	An attractive, abstract representation of sound volume and pitch, that moves around screen to indicate sound location (see Figure 2b).
		👎	You Map	‘You’ are always at the centre of the screen and sounds are displayed as coloured rings relative to you, indicating their location.
	Small	👉	Rings	Coloured rings appear at edges of a PC monitor, indicating volume (ring size), pitch (colour), and relative location (<i>e.g.</i> , rings on the left side of screen occurred to the left of the screen, <i>etc.</i>). (See Figure 3a).
		👉	LED Panels	A row of physical LEDs on either side of a PC or TV screen, indicate volume (number of lit LEDs), pitch (colour), and relative location (same as Rings).
		👉	Colour Border	A border around PC screen filled with colour indicating relative locations of sounds (same as Rings), volume (colour intensity), and pitch (colour chosen).
Proximity, volume, pitch	Small	👎	Location Bubbles Sidebar	Sidebar on PC screen with a row of bubbles that indicate proximity of sounds to screen with bubble size (<i>e.g.</i> , if sound is close to screen, bubbles higher on sidebar are larger, if sound is far from screen, lower bubbles are larger).
Any	Small	👍	PDA Display	Any of the other displays adapted to a PDA. Vibrates for alerts.

Figure 3

Figure 4

Figure 5

Figure 6

Table 3

Prototype	Position	Amplitude	Frequency	Sound Location	Sound Recognition	Training	Wizard of Oz	Customize	Confidence	History
Single Icon	No	No (but, did show whether hi or lo)	No (but, did show whether hi or lo)	No	Yes	Yes	No	Yes	Yes	Yes
Spectrograph with Icon	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Spectrograph	No	Yes	Yes	No	No	No	No	No	No	No
Map	Yes	Yes	No	Yes	No	No	Yes (sound location)	No	No	No

Figure 7

Figure 8

Figure 9

Figure and Table Captions:

- Figure 1: (a) Implemented version of *Spectrograph with Icon* showing a phone ringing. (b) *Single Icon* showing a loud unrecognized sound (many rings indicates a high volume) with low frequency (indicated by the blue colour of rings). Image originally published in [13], Copyright ACM Press, 2005.
- Table 1: Taxonomy of existing non-speech sound awareness techniques for the deaf. Y means a technique supports at least some examples of that dimension. H, M, and L refer to High, Medium and Low, respectively.
- Figure 2: (a) Map showing locations of sounds with rings. (b) Ambient Visualization showing a range of volumes, from none (top left) to high/loud (bottom right). Image originally published in [13], Copyright ACM Press, 2005.
- Table 2: Design sketch descriptions. Thumbs indicate overall user opinions: up = like; down = dislike; sideways = mixed.
- Figure 3: (a) Rings, showing relative locations of sounds: a loud sound to the left and behind the screen, and a soft sound to the right. (b) Directional Icons, showing a voice to the left, a phone ring in front, a door knock to the right, and a very loud, unrecognized sound behind the screen. Image originally published in [13], Copyright ACM Press, 2005.
- Figure 4: History Display shows each sound recognized as a different coloured bar along a time axis (letters have been added here for clarity: p = phone, d = door, s = speech). The bar height indicates the sound’s relative volume. Image originally published in [13], Copyright ACM Press, 2005.
- Figure 5: Speech visualized by the Spectrograph (without icon) prototype. In this visualization, height is mapped to frequency, colour to intensity (blue = quiet; green = medium; red = loud). The temporal aspect is depicted by having the visualization animate from right to left. Image originally published in [8], Copyright ACM Press, 2005.
- Figure 6: Speech visualized by the Map prototype. The speech is coming from the second desk in the room and is identifiable by its position (left of the map). Both the number of rings and the colour of the rings indicate the volume of the sound (together, they provide redundancy).
- Table 3: Comparison of our four implemented prototypes.
- Figure 7: Correct identifications: (left) Average correct identifications for each display in Quiet and Noisy conditions. (right) Certainty of correct identification (5 is most certain) for each display in Quiet and Noisy conditions.
- Figure 8: False alarms: interaction effect of Display, Signal, and Environment on false alarms (y-axis). Error bars show 95% confidence intervals of the mean. Dots and lines show means.
- Figure 9: Distraction: a comparison of the number of zeros that were detected per second in Quiet and Noisy conditions, for each display. Fewer zeros indicates greater distraction. Error bars show 95% confidence intervals of the mean. Dots and lines show means.