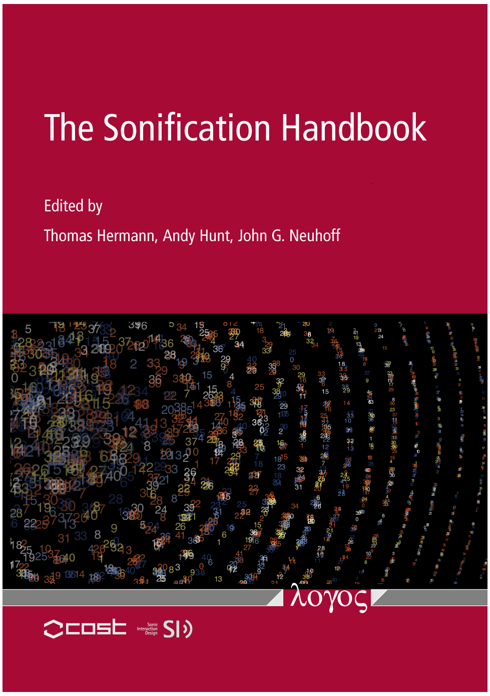

The Sonification Handbook

Edited by
Thomas Hermann, Andy Hunt, John G. Neuhoff

Logos Publishing House, Berlin, Germany
ISBN 978-3-8325-2819-5
2011, 586 pages

Online: http://sonification.de/handbook
Order: http://www.logos-verlag.com

Reference:
Hermann, T., Hunt, A., Neuhoff, J. G., editors (2011). *The Sonification Handbook*. Logos Publishing House, Berlin, Germany.

Chapter 17

Auditory Display in Assistive Technology

Alistair D. N. Edwards

Auditory information can be of particular importance to people who cannot perceive other forms, notably those who are blind. This chapter mainly surveys some of the attempts to substitute visual information by sounds.

Reference:
Edwards, A. D. N. (2011). Auditory display in assistive technology. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 17, pages 431–453. Logos Publishing House, Berlin, Germany.

Media examples: http://sonification.de/handbook/chapters/chapter17

Chapter 17

Auditory Display in Assistive Technology

Alistair D. N. Edwards

17.1 Introduction

This chapter is concerned with disabled people[1]. As soon as a label such as 'disabled' is applied, questions are raised as to its definition. For the purposes of this chapter, no formal definition is required, rather it should suffice to say that the people we are writing about have the same needs as everyone else, it is just that in some instances their needs are more intense and are sometimes harder to meet. If this book achieves anything, it should convince the reader that sound can be an immensely powerful medium of communication and the relevance of this chapter is that the full potential of the use of sounds can often be more completely realized when aimed at meeting the needs of people with disabilities.

The immediately obvious use of sounds is as a replacement for other forms of communication when they are not available. Specifically, blind people cannot access visual information. Much of this chapter will deal with this form of substitution, but it will also demonstrate the use of sounds in other applications.

It is a contention in this chapter that there is a great potential for the use of sound that has not yet been realized, but some progress has been made in the following areas which are reviewed in this chapter:

- computer access

[1] Language is powerful and sensitive. No other literature is more sensitive to the needs of being politically correct than that which deals with disability. It is recognized that inappropriate use of language can cause harm and offence, but at the same time perceptions of what is correct are constantly changing. For example, at the time of writing there are (sometimes fierce) arguments as to whether 'disabled people' or 'people with disabilities' is the better term. In this chapter we have attempted to be sensitive to all shades of opinion, and if we have failed and used any terminology felt to be inappropriate by any individual reader, then we can only apologize.

- mobility aids.

Then there are other potential uses and some of these are also discussed.

17.2 The Power of Sound

Of course, one of the most powerful (and the most used) form of auditory communication is speech. Even though the emphasis of this book is on non-speech sounds, the role of speech cannot be ignored and it will be discussed in this chapter, in the context of where speech has advantages over non-speech.

The potential power of non-speech sound is illustrated by the following extract, written by John Hull, who is blind.

> I hear the rain pattering on the roof above me, dripping down the walls to my left and right, splashing from the drainpipe at ground level on my left, while further over to the left there is a lighter patch as the rain falls almost inaudibly upon a large leafy shrub. On the right, it is drumming with a deeper, steadier sound, upon the lawn. I can even make out the contours of the lawn, which rises to the right in a little hill. The sound of the rain is different and shapes out the curvature for me. Still further to the right, I hear the rain sounding upon the fence which divides our property from that next door. In front, the contours of the path and the steps are marked out, right down to the garden gate. Here the rain is striking the concrete, here it is splashing into the shallow pools which have already formed. Here and there is a light cascade as it drips from step to step. The sound on the path is quite different from the sound of the rain drumming into the lawn on the right, and this is different again from the blanketed, heavy, sodden feel of the large bush on the left. Further out, the sounds are less detailed. I can hear the rain falling on the road, and the swish of the cars that pass up and down. I can hear the rushing of the water in the flooded gutter on the edge of the road. The whole scene is much more differentiated than I have been able to describe, because everywhere are little breaks in the patterns, obstructions, projections, where some slight interruption or difference of texture or of echo gives an additional detail or dimension to the scene. Over the whole thing, like light falling upon a landscape, is the gentle background patter gathered up into one continuous murmur of rain. [1, p. 26-27][2]

There are two important points to be taken from this extract. Firstly there is the immense amount of information that the writer was able to extract from sounds. Secondly, it has to be acknowledged that none of the attempts to use sounds in synthetic auditory displays has yet come close to conveying that amount of information. It can be done; we do not yet know how to do it. It has to be acknowledged that most of the devices and ideas described in this chapter are *not* embodied in commercially available, commonly-used products. For various reasons they are not sufficiently useful for widespread adoption, and yet the above extract clearly demonstrates the richness of information that can be usefully conveyed in non-speech sounds. Tony Stockman also describes how blind people can make use of environmental sounds, putting them in the context of attempts to supplement these with technology-generated

[2] *On Sight and Insight*, © John M. Hull, 1990, 1997. Reproduced by permission of Oneworld Publications.

sounds in [2].

Sight is a very powerful sense. By any measure, the amount of information that can be received visually is vast. Yet it is not simply the raw bandwidth of sight that makes it so powerful, it is the ability to (literally) focus on the information that is of relevance at any given time. Because of the amount of information available visually it is those people who do not have access to visual information who are the most obvious candidates to use auditory information as an alternative. There is a fundamental problem, though, in substituting for visual information. The capacity of the non-visual senses (including hearing) simply does not match that of sight. This is often referred to as the *bandwidth problem*.

Thus, the fundamental restriction is that sounds cannot be used to convey the same amount of parallel information as the visual sense can. There are two principal approaches that can be taken to address this problem:

1. Maximize the amount of information carried in the sounds;
2. Reduce the amount of information presented (i.e. filter it in some way).

Achieving (2) amounts to giving users a form of focus control corresponding to that of the visual sense. While (1) is the main topic of this chapter, it cannot be divorced from the necessity to provide the control implied in (2).

In this chapter a number of research projects are described in which non-speech sounds are used to convey information to blind people. In comparison to the example from John Hull, above, it will be evident that these attempts are quite crude. Nevertheless, this is surely a stage that has to be gone through in order to understand the nature of this style of communication, with the hope that eventually we will be able to create vast, rich and useable soundscapes.

17.3 Visually Disabled People

There are a large number of people with visual impairments. Although exact figures are hard to find, Tiresias [3] estimate that there are approximately six million people in Europe with a visual disability. Visual disabilities take a number of forms and the number of blind people - those with no useful sight - is relatively small (one million in Europe, according to Tiresias)[3].

Although the number with an impairment short of blindness (variously referred to as ‘visually impaired’ or ‘partially sighted’) is relatively large, the number of different forms of impairment make it difficult to meet their needs. (An impression of the effects of different forms of impairment can also be found on the Tiresias website, [3]). For instance, an adjustment that helps some people (such as text enlargement for people with cataracts) can even make vision worse for others (enlargement further reduces the material in view to someone with tunnel vision, perhaps due to glaucoma).

It might be suggested that any interaction that makes no use of vision - such as an auditory

[3]The figures are open to debate. For instance, [4] estimated the number of visually disabled Europeans as 2,000,000, while the proportion of people with visual disabilities has been estimated variously as 1.6% (in Europe, [3]) and 4.1% (in the USA, [5]). Also, one must beware that the population of Europe has changed since 1993 with the accession of new states.

interface designed for users who are completely blind - would be equally accessible to those with some vision. However, the fact is that those with some sight generally prefer to make as much use of that sight as possible. In other words, they do not need substitution of the visual information, but its enhancement to match their visual abilities.

Thus, this chapter really addresses the needs of those who must have an auditory substitute for visual forms of information, those who are blind - even though they are the minority of those with visual disabilities.

Visually disabled people have a variety of needs for non-visual information. This chapter looks at access to computers (through screen readers), electronic travel aids and other applications which make use of sounds.

17.4 Computer Access

Most human-computer interfaces are 'visually dominated' in that the principal channel for communication from the computer to its user is the monitor screen. For a blind user, all the information that is displayed on a computer screen has to be substituted by non-visual forms of communication, either tactual or auditory.

The dominant form of tactual communication is braille. Braille is mainly a translation of printable text. The greatest barrier to the use of braille, though, is the small number of (blind) people who have the skills to read it. Again accurate statistics are hard to compile, but Bruce at al. [6] suggest that in the UK the proportion of blind people who can read braille is as low as 2%. Computer braille displays are available [7]. These usually consist of 40 or 80 braille cells. They are electro-mechanical devices and are thus quite expensive and are also bulky and heavy.

While there is a significant community of enthusiastic braille users - including those who use braille for computer access - auditory interfaces have a lot of features which make them very attractive compared to braille, notably:

Ease-of-use: Unlike braille, sounds essentially require no training. Of course this is not strictly true of some of the more complex uses of sounds discussed in this book (e.g. chapters 8, 10, 12, 14), but the simplest sounds including speech can be used without training. Auditory interfaces are effectively accessible to 100% of blind people - as long as they do not also have a hearing impairment.

Cost: Sound cards are a standard component of all modern PCs, therefore the only additional cost is that of any special software.

Braille was originally designed for the presentation of literary text - that which can be expressed in the 26 letters of the alphabet plus 10 digits and a small number of punctuation marks. Its extension to other forms of communication (e.g., mathematics or music) is somewhat clumsy and labored. There is a similar problem with sound when applied to the complex information that can be displayed on a computer screen. On any computer screen there may be hundreds of different elements visible. The sighted user can cope with this large amount of information because they have the ability (literally) to focus on the item of interest at any time. Thus, the user can take in the information of importance and filter out that which is currently irrelevant. The non-visual senses (and here we are mainly concerned with hearing) do not have that ability.

In other words, if we were to take a simple-minded approach to the adaptation of a visual display for blind users we might try to associate a sound with each item on the screen. To glance at such a screen would imply having every one of those items make its sound. Clearly this would be a cacophony. Sounds would interact and mask each other and it would not be possible to spatially separate the sounds in the same way that vision can do.

Computer access for blind people is achieved by using a piece of software called a *screen reader*. Essentially this examines the contents of the screen and converts it into sounds[4].

Screen readers were first developed about the same time as the PC became available. The operating systems of the time (predominantly MS-DOS) were text-based. That is to say that the screen displayed text and commands were typed in on the keyboard. For instance, to display the contents of the current directory, the user would type `DIR`, or the contents of the file `foo.txt` could be displayed (*typed*) on the screen by entering `TYPE FOO.TXT`. It was relatively easy to render this kind of interaction (i.e., the text of the command line and the contents of the text file displayed in response to the command) in sounds by using a screen reader linked to a speech synthesizer. Some of these first-generation screen readers made some use of non-speech sounds. For example, the *Hal* screen reader [9] used beeps of different tones to guide the user between the different lines on the screen, but most of these screen readers relied mainly on speech.

The screen reader represented a major advance for blind people. The access to the computer that it gave, generated a degree of equality in job opportunities; jobs that had been inaccessible now became feasible for blind workers.

The next major development in the personal computer was the graphical user interface (GUI). This was firstly implemented commercially on the Apple Macintosh, but eventually was also found on 'IBM-compatible' PCs in the form of MicrosoftWindows. At first the GUI was seen as a real threat to blind people. The form of interaction was completely different and very much visually orientated. The mouse pointing device was added to the keyboard and screen. It was necessary to point at objects on the screen. The design and positions of those objects carried meaning. These properties and their meanings could not easily be translated into auditory forms. The emancipation that blind workers had experienced was in danger of being lost.

Edwards [10, 11] experimented with an auditory version of the GUI, *Soundtrack*. This was not a screen reader, but a word processor which retained most of the interactions of the GUI (windows, icons, scrollbars etc.) but represented them in an auditory form. The first level of interaction was based on tones of varying pitch, giving relative spatial information, but at any time the user could click the mouse and hear a spoken label. Double-clicking would activate the current object.

Soundtrack remains one of the few attempts to make mouse-based interaction with a GUI accessible in a non-visual form, but it was only a word processor, and not a generalized tool for making GUIs accessible. However, screen readers were eventually developed such that the modern GUI interface is about as accessible as the former text-based ones were. GUI screen readers obviate the need to use the mouse by taking over control of the cursor, which

[4] Most screen readers can also render the information on a braille display [7, 8], but that is outside the scope of this chapter.

is controlled through the keyboard. They also still tend to rely to a large extent on synthetic speech with minimal use of non-speech sounds.

Syntha-Voice's *WindowsBridge* was a screen reader which attempted to make the Windows operating system accessible through the mouse. Positional feedback on the cursor was given using musical tones, and mouse movements could be filtered so that only vertical and horizontal movements were detected (i.e. no diagonal movements). However, few users used this feature and, indeed, the product is no longer available.

Non-speech sounds were used more extensively in some experimental screen readers, notably *Mercator* and *Guib*. The contrasting approaches behind these different systems is written up in [12], but both tended to use the style of non-speech sound known as the *auditory icon* [13] (chapter 13). The Guib Project culminated in a commercial screenreader, *Windots*, but it did not make much use of the non-speech sounds developed in Guib.

Windots was never a great success commercially and is no longer available. In practice the most popular Windows screen reader is *Jaws for Windows*[5]. Jaws has quite extensive facilities for the use of non-speech sounds. A Speech and Sound Manager allows users to associate different utterances or sounds with screen objects. These include:

Control types These are widgets, such as buttons, scrollbars, check boxes.

Control state Widgets can be rendered differently depending on their state, a button that is pressed or a check box that is checked or not.

Attributes Different font attributes can be signaled.

Font name Changes in font can be signaled.

Color The color of the current item can be signaled.

Indentation An indication of the depth of indentation is presented.

HTML Different HTML elements (in webpages) can be signaled.

All of these properties can be rendered in different ways. Speech may be used (i.e., an explicit description of the attribute) or a change in the current voice, but there is also the option of playing a sound. Sounds are simply played from *.wav* files and a number of these are provided with the Jaws software. These include auditory-icon-style sounds such as recordings of door bolts being opened or closed (sample **S17.1**), a lamp being switched on (sample **S17.2**), the thump of a rubber mallet (sample **S17.3**) and the like. There are also musical sounds, such as a piano playing one or two notes (e.g., sample **S17.4**) that can be used in a more earcon-style of soundscape.

Different 'Speech and Sound' configurations can be created and stored. This means that users can load particular configurations for different purposes. For instance, they may wish to use one configuration when word processing and a different one when writing programs. Configurations can be stored in files. This means that they can easily be swapped and shared between users. A number of configurations are provided with the Jaws software and it is interesting that these make minimal use of the sounds option; they are (again) speech-driven.

Microsoft Windows is the operating system of choice for most blind users; it is best supported

[5]Freedom Scientific (`http://www.freedomscientific.com`).

by available screen readers. The Unix world has been slow to recognize the needs of blind users, but this changed with the advent of the Linux Gnopernicus Project which aimed to enable users with limited vision, or no vision, to use the Gnome 2 desktop and applications effectively. However, this project appears to have stalled due to lack of funding.

As mentioned above, the advent of the GUI was seen at the time as a serious blow to the emancipation of blind computer users. Apple Computers were responsible for the introduction of the GUI to the consumer market, with the release of the Macintosh. Although a screen reader (*OutSpoken*) was released for the Macintosh, it has never been used by many blind users (and is no longer marketed). However, in the release of Version 10.4 of its OS X operating system, Apple included *VoiceOver*, a built-in screen reading facility. Controversy exists regarding the efficacy of this screen reader [14, 15], but its use is growing as it is now part of the iPhone and the iPad. As with most screen readers, it is heavily speech-based, but does include the use of non-speech sounds.

A common theme in this book is that the true potential for the use of non-speech sounds has yet to be realized. This is clearly true in the application of computer access for blind individuals. Most screen readers have facilities for the use of non-speech sounds; however, few people use them. This implies that the kinds of sounds being used and the information they are providing is not perceived as valuable to the users.

17.5 Electronic Travel Aids

The need to access computers is growing, but is still a relative minority activity compared to moving around the world. There are two aspects to this for blind people: short-range obstacle avoidance and the broader-scale of navigating to desired destinations. Technologies (sometimes referred to as Electronic Travel Aids or ETAs) can be used in both of these applications.

By far the most popular technology for obstacle-avoidance is the guide cane (also known as the white cane). There are a number of reasons why this is so popular, which will be discussed in contrast to higher-technology approaches below. For a person walking through an environment, it is vital to know whether there are any obstacles in the path ahead. This is what the guide cane provides. Canes come in different lengths from approximately 60cm to 160cm. Shorter canes are symbolic-only, carried by the user (who is likely to have some vision) as a signal to others that they may need special assistance. It is only the longer ones that are used for obstacle avoidance.

The cane communicates information mainly through the haptic senses. In other words, the user detects forces on the cane handle as its tip collides with objects. However, it is important to be aware that there is an auditory component to the communication also. The sound that the cane makes as the tip is tapped on surfaces can communicate a lot of information. For instance the texture of the path (e.g., concrete versus grass) will be apparent from the sound the tip makes. Also the sounds generated will be modulated by the environment. A closed area surrounded by walls will generate echoes, whereas an open one does not. The amount of information available from such natural auditory sources should not be underestimated. John Hull describes [16] how he can recognize when he is walking by railings by the intermittent echo that they generate[6]. Snow is sometimes described as ‘the blind person’s fog’ - because it dampens sounds rather as fog blocks sighted people’s vision.

One of the major disadvantages of the traditional cane is that it operates only within a very narrow vertical range. That is to say that it will generally detect obstacles at ground level. While that is sufficient in many environments, there is clearly a danger from any obstacles up to head height. This is one advantage that high-technology sensors can have; they can scan the entire path ahead of the user.

There is then the question as to how to communicate the information to the user, in a non-visual form. Sound is the obvious medium to use [18]. There are, however, two particular problems with sound: auditory interference and the bandwidth problem.

A question arises as to how to present the auditory feedback from a guidance device. Headphones may seem the obvious choice. They can present the information privately. This is important because the information is not of any use to anyone else in the vicinity and is therefore likely to annoy them. More importantly, any audible sounds would draw attention to the person generating them and might be perceived as a label of their blindness. Headphones can also be used to present information spatially, either using simple binaural stereo or three-dimensional spatializations. Finally since the advent of the portable stereo player, it has become socially acceptable to wear headphones in public, so that their use is not a social faux pas.

However, headphones are in practice not necessarily appropriate. The main problem is that they are tend to interfere with environmental sounds. There are various headphones available designed to avoid this problem by not blocking external sounds [19], but their effectiveness is open to question.

Conspicuity and aesthetics are important factors, the importance of which are easy to underestimate. Most people do not like to stand out in the crowd, and this is just as true of people with visual disabilities as for sighted people. Modern white canes are usually foldable. That is to say that they can be dismantled and folded into a package around 20cm in length. This means that their visibility is under the user’s control. As illustrated by the symbol cane, one of the features of the white cane is that it can be positively used as a signal to other people that the owner has a visual disability. However, on the other hand, the user can also choose to fold the cane away, removing that signal.

Some high-technology devices are not so discreet. For instance, the *Kaspa* [20] is a box worn on the forehead. While improvements in miniaturization will almost undoubtedly make it possible to conceal such devices better, most users will still not want to wear equipment that is too visible. Any such device makes the user stand out, clearly indicates that there is something different about them (it may not be obvious that the person has a visual disability) and may make them seem to be quite odd and freakish.

The importance of aesthetics should also not be underestimated; even a device which is very positive in the assistance and power that it gives the user, will be rejected by many if it is too ugly. (See also Chapter 7).

There has already been mention above of changing attitudes towards headphones, which

[6] Another interesting example is the experiments by McGrath et al. [17] in which it was found that blind people could locate and accurately describe objects (a sheet of aluminium, a sheet of aeroboard and a leather football) in a dark room using only the sound of their voice.

also suggests a fashion element. With the advent of the Sony walkman in the 1980s, it became acceptable (at least among young people) to wear headphones in public. For the most part they were small, discreet, and not very noticeable. Now the wearing of headphones is common. In fact, in some environments (such as on public transport) there may be as many people wearing headphones or earphones as not. Yet fashion has also moved on. There are those now who prefer not to wear barely visible earphones, but rather large, highly conspicuous headphones. They would, no doubt argue that their choice is based on acoustics, that the sound reproduction is so much better, but at the same time, the headphones are usually high quality ones - of sleek design and with the accompanying clear brand labels. In other words, there is an element of boasting in the wearing of these devices.

It would be ideal if the same kind of positive kudos could be attached to aids for visually disabled people. In other words, the device could become something 'cool' and not a label of deficiency. The headphones example illustrates, though, that aesthetics and fashion can involve complex interactions.

Yen [21] provides a comprehensive list of ETAs, some of which are explored in more detail in the following sections. While the emphasis is on the technical specification of these devices, it should be apparent that other factors - including aesthetics - are also important.

17.5.1 Obstacle Avoidance

There are a number of devices which operate as obstacle detectors. It is significant that the same approach to obstacle avoidance - a portable sensor generating auditory signals - has been tried many times. There is no point in trying to provide an exhaustive list of such experiments, but several devices are reviewed in [22], and [23] including:

- Russell Pathsounder,
- Nottingham Obstacle Detector,
- Laser Cane,
- Sonic Torch,
- Mowat Sensor,
- Sona,
- NavBelt.

Two exceptional examples are the *Bat 'K' Sonar-Cane*[7] and the *UltraCane*[8], exceptional in that they are commercially available products. The 'K' Sonar is a hand-held device resembling a flashlight or torch that can be clipped to a white cane. It has a cable connection to two miniature earpieces. The pitch of the echo sounds is proportional to distance: high-pitched sounds relate to distant objects and low-pitched sounds related to near objects. (Examples are provided, see Table 17.1).

One feature of many of these hand-held devices is that they are directional. That is to say that they generate feedback about obstacles only when they are within the (narrow) beam of the device. This means that spatial information is given directly by the device (i.e., through the

[7] http://www.batforblind.co.nz
[8] http://www.ultracane.com

Description	File name
Scanning a 4cm-diameter plastic pole at 1.5m from left to right and back.	**S17.5**
Walking towards a glass door from a distance of 5m to 1.5m and then retracing steps back to 5m. Person approaching from 5m to the halt position and then retracing his steps.	**S17.6** **S17.7**
Standing in front of a wooden fence with spacing between small panels and scanning the torch to the left and right to 'shine' the beam along the fence line.	**S17.8**
Scanning the torch down onto a grass lawn and up again.	**S17.9**
Standing in front of a 50cm wide tree. The tree has only a thin layer of bark and a thin (4cm) shoot growing out at the base of the tree; a clump of short flax on the other side. As the ultrasonic beam scans across it produces a strong 'warbling' sound from the trunk, a soft mushy sound from the flax, and a soft short whistle from the shoot.	**S17.10**
Standing in front of a concrete block wall with large thick well-developed ferns at the side of the standing position. The torch scanned across the ferns onto the wall and back again. The wall produces a tone sound. The ferns made a strong mushy sound.	**S17.11**

Table 17.1: Sample K-Sonar sounds.

kinesthetic information that the user has about the position of the hand grasping the device). There is no requirement to encode spatial (directional) information in the auditory signal. This makes the signal simpler - and hence generally easier to comprehend. In other words a hand-held device sends out a one-dimensional beam and scanning it horizontally adds a second dimension of information, whereas a representation of the entire scene includes all three dimensions. These might be represented directly by spatialization of the auditory representation (as in the *Kaspa*, [20]) or by applying some other modulation to the signal.

The UltraCane is also important in that it is commercially available. It avoids the problems of auditory output discussed above by presenting its information haptically. As such, it is outside the scope of this book, but it is interesting in that it possibly illustrates an attempt to sidestep some of the disadvantages of using sound. Sounds - and headphones - are not used, so there is no masking of the natural acoustic environment. The mapping of obstacles to vibrations of different buttons in the cane handle, with strength indicating separation, is a natural one.

As mentioned earlier, despite the advent of clever electronic aids, the white cane remains the most popular device. It is worthwhile looking at reasons for this. Differences between guide canes and electronic alternatives are summarized in Table 17.2.

The NavBelt is an experimental device of interest because of the ways it operates [23, 24, 25, 26, 27]. It takes the form of a belt worn around the user's waist. The belt contains an array of sonar devices. The sonars measure the distance to obstacles. NavBelt operates in two modes. In Guidance Mode it is designed to *actively guide* the user around obstacles towards a target, while in Image Mode it substitutes an auditory scene for the visual scene (more akin to the kinds of visual substitution systems explored in the next section).

The sonars detect obstacles from which the NavBelt calculates the *polar obstacle density*, a measure which combines the size of obstacles and distance to them. In Guidance Mode, the NavBelt calculates the area with the lowest polar obstacle density near the direction of travel and guides the user in that direction. In other words, Guidance Mode works best when the target is known. This might be achieved by integrating the NavBelt with a navigation aid. Bornstein [24] lists this as a potential future development, but there is no evidence of this having been subsequently implemented. In the absence of an absolute means of specifying the target, the device uses heuristic approaches to infer the intended direction of travel.

Cane	**Electronic obstacle detector**
Inexpensive. Losing one or accidentally swapping with another owner is not a major problem. It is feasible to own more than one in case of loss or damage. A standard guide cane costs of the order of €20 or $30.	Expensive. The ‘K’ Sonar costs of the order of €500 or $650, and the Ultracane is around €750 or $900.
Reliable.	Subject to faults and requiring maintenance.
Does not interfere with hearing.	Acoustic signals may block natural cues.
Senses only at ground level.	Can be designed to sense up to head height, but may not detect some important ground-level obstacles (e.g., kerbs).
Extensive training required - over 100 hours.	Estimates and claims as to the amount of training required vary.
Short-range - effectively the length of the cane (1 - 2 meters).	Can be designed to operate at longer ranges. Typical sonar devices can operate up to 10 meters. Video-based systems theoretically can operate up to the visual horizon.
Requires constant active exploration.	Requires constant active exploration.

Table 17.2: Comparison between the features of the traditional guide cane and electronic alternatives.

It is interesting that technological orientation devices have been under development for at least thirty years (e.g.. [28]). The ‘K’ Sonar had achieved sales of 1550 up to 2010[9] - but this is a tiny proportion of the market.

The guidance information is presented to the user as binaural sounds on headphones with interaural time difference to create the impression of directionality. In Guidance Mode the pitch and amplitude of the sounds are proportional to the recommended travel speed. The principle is that higher pitch and amplitude attract attention so that the user will instinctively slow down and concentrate on the direction of the signal. A special low-pitch signal (250 Hz, near to middle C) is generated when the direction of motion is approximately correct (i.e.,

[9]Personal communication

Sighted		1.30
Image Mode	Simulation	0.52
	Physical	0.40
Guidance Mode	Simulation	0.76
	Physical	0.45

Table 17.3: Average walking speeds (ms^{-1}) under different conditions. 'Sighted' refers to the speed of the average sighted walker. The other figures relate to the evaluation of the NavBelt in its two modes, both in simulations and in physical traversal of a laboratory. Note that the figure for Image Mode (Physical) was only attained after 'several hours' of training.

within $\pm 5°$). This provides simple positive feedback when the user is going in the correct direction. At the same time, using a low-frequency tone will have less of a masking effect on environmental sounds.

Image Mode is designed to invoke the impression of a virtual sound source sweeping across 180° in front of the user. The sweep is completed in 37 discrete steps separated by 5°. The sounds used are square waves modulated by amplitude, pitch and duration [26]. The duration of a signal varies between 20 and 40ms, where 20ms indicates the longest distance to an obstacle (5 meters) and 40ms indicates a very close object (0.5m). The amplitude varies inversely with the range reading from the corresponding sonar sector. Sixteen discrete amplitudes can be selected, where the lowest value (silence) represents no threat to the user from that direction, whereas the maximum value indicates a high risk. The intention is that 'the user's mind creates a mental picture of the environment that adequately describes the *obstacle density* around the user.' [24, p. 113].

Evaluations of the NavBelt have been based on simulations. In navigating randomly selected (simulated) maps the average travelling speed was $0.52ms^{-1}$ (compared to an average sighted person's walking pace of around $1.3ms^{-1}$). It was evident that the walking speed depends very much on the complexity of the environment. A more complex environment requires greater cognitive effort by the user and apparently leads to a slower walking speed. At the same time there appeared to be a learning effect, whereby experienced NavBelt users attained higher speeds. It was also noted, though, that users with 'reduced auditory perception capabilities travel slower than highly skilled people.' [26]. As well as the simulations, experiments were also carried out using the actual NavBelt in which blindfolded participants travelled from one side of the laboratory to the other. Walking speeds were slower than in the simulation because participants were more cautious. 'However, after a training period of several hours they traveled safely through the controlled environment of the laboratory with an average speed of $0.4ms^{-1}$.' (*ibid.*)

Simulation evaluations of Guidance Mode showed an average travel speed of $0.76ms^{-1}$ and an average deviation from the recommended direction of 7.7°. In similar experiments with the real prototype NavBelt whereby participants travelled 12 meters across the laboratory the average speed was $0.45ms^{-1}$. A more-realistic experiment was carried out in an office building corridor with which the participants were familiar. The length of the path was 25 meters and included several corners. No obstacles were positioned initially in the corridor, but passers-by did walk down the corridor. Participants were also able to avoid obstacles and attained an average walking speed of 0.6 ms^{-1}. The walking speeds attained under the various conditions are summarized in Table 17.3.

17.5.2 Visual Substitution

Developers of the NavBelt have experimented with two approaches to guidance. Its Image Mode is an example of the approach whereby the idea is to generate an auditory field representing the entire visual scene that a sighted person would see. The visual picture can be captured through video cameras and then translated into an auditory form.

The bandwidth problem was described above. The same problem applies in this application. Sensors such as video cameras can provide large amounts of (visual) information. The question is how much of that information to provide to the user. The more information, the greater the user's freedom to navigate, but the harder it becomes to interpret and understand. Obstacle detectors, such as those described above, can generate quite simple sounds, giving an indication of the location of the size and location of objects. At the other end of the scale there have been attempts to render the entire scene sonically.

One example is the *Voice* Project (see [29, 30] and the Voice web site[10]). This creates a representation of the visual picture pixel-by-pixel. The vertical positions of pixels are represented by pitch, horizontal positions (left-to-right) are represented by time, and brightness is represented by loudness. The sound effectively scans horizontally across the image so that a vertical column of pixels are all presented in a single, complex sound. The start of a scan is marked by a 'click' and the scanning repeatedly loops. An example of this sonification is shown in Figure 17.1.

This simple mapping is quite raw, implying minimal processing of the image. The system relies instead on brain plasticity. The intention is that with practice the user will learn to interpret the auditory scenes naturally. Some support for this approach is given in [31] which describes the examination of functional magnetic resonance images (fMRI) of the brains of blind and sighted participants performing sound localization tasks. They observe that blind people demonstrate a shift in activated brain areas towards more posterior areas - the areas that are involved in visual processing in sighted people[11].

González-Mora et al. [31, 37] have experimented with a prototype device which incorporates video cameras and headphones mounted on a pair of spectacles. Their sonification is described as follows:

> 'The basic idea of this prototype can be intuitively imagined as trying to emulate, using virtual reality techniques, the continuous stream of information flowing to the brain through the eyes, coming from the objects which define the surrounding space, and which is carried by the light which illuminates the environment. In

[10] http://www.visualprosthesis.com/voice.htm

[11] Modern brain imaging techniques such as fMRI have enabled researchers to shed new light on the idea that blind people's non-visual senses are in some ways heightened. Previously there was some skepticism about this apparent phenomenon (e.g., [32]) but now there is an increasing body of knowledge which suggests that the area of the brain usually referred to as the *visual* cortex is devoted largely to the processing of visual information simply because in sighted people that is the predominant source of stimulation. In people deprived of sight, the same area can be reassigned to the processing of non-visual information. Examples of this work include [33, 34, 35, 36].

Figure 17.1: Sample graphic which is sonified by the Voice system as illustrated in Sample **S17.12**. Note that the blurred style of the picture is deliberate, reflecting the pixel-by-pixel translation to sound.

> this scheme two slightly different images of the environment are formed on the retina, with the light reflected by surrounding objects and processed by the brain to generate its perception. The proposed analogy consists of simulating the sounds that all objects in the surrounding space would generate' [31, p. 371-372].

Of course in a real environment, inactive objects do not generate sounds, but in the prototype system a click sound is used:

> 'When a person is in front of a particular scene, he/she receives an acoustic input consisting of a set of a set of auralized[12] "clicks", with a randomized order of emission, corresponding to the calculated 3-D coordinates of the objects. This set of "clicks" is sent to the person in a time period of 153ms, after which the next acoustic image is sent. Depending on the number of coordinates that the objects occupy inside the perception field, there is a variable interclick interval, never less than 1 ms.' (*ibid*. p. 374).

Interestingly, in the context of the earlier quote from John Hull, the perceived effect is described as resembling 'a large number of rain drops striking the surface of a pane of glass'.

Spatial information is reproduced by spatialization of the sounds, using individualized head-related transfer functions (HRTFs)[13]. A field of 80° horizontally by 45° vertically is presented with a resolution of 17×9 and 8 levels of depth (although higher resolutions are being developed).

[12] The authors appear to use the word 'auralized' to mean 'spatialized'.

[13] Every individual is different in the way they perceive spatial sounds because of the shape of their ears and head. This can be modeled for artificially spatialized sounds by creating their HRTF. Best results are achieved using individual HRTFs, although an 'average' HRTF can be used, but it will be less effective.

Evaluations have yielded results which are claimed to be encouraging regarding blind people's ability to perceive the layout of a test room - although the evaluation is not described in detail. It is interesting that some participants reported experiencing apparent synaesthesic effects, whereby the sonification evoked a visual perception of 'luminous sparkles' coinciding with the spatial location of the sound sources. The system is very much a prototype and not yet a released product.

Sighted people rely on light reflecting off objects in the environment entering their eyes and forming an image on the retina. Babies learn to interpret these images through active exploration of their environment and hence learn to rely on them in interacting with the world. Visual substitution systems aim to create a similar representation using sounds. Most objects do not make sound, though, so there is no natural acoustic 'light', so instead a visual image is translated by technology into sound. The hope is that people can learn to interpret these soundscapes as well and as naturally as visual scenes. There is some hope for this approach in that the plasticity of the brain in interpreting acoustic input has been well demonstrated by the success of cochlear implants for deaf people. A cochlear implant generates artificial sensations in the auditory nerves. There is no reason to believe that the sensations thus generated resemble those generated by natural hearing, and yet - with practice - people become quite adept at interpreting those inputs as if they are (low fidelity) sounds [38].

It has to be stated that there is a dearth of formal evaluations of most of the systems described in this section. This is clearly a weakness in research terms, but furthermore there must always be a fear that the systems are ineffective and that to pursue them further would be a waste of time.

17.5.3 Navigation Systems

The obstacle avoidance and visual substitution systems described above are predominantly concerned with short-range mobility, mainly the avoidance of obstacles. A different problem is that of navigating to a chosen destination. For instance, when a person arrives by train in a strange city, they may need to know how to get to an office block which is known to be walking distance from the station. This is a problem for all travelers, but sighted people can rely on maps and similar aids to work out and follow the correct route. It is increasingly common now for car drivers faced with such navigation problems to rely on a SatNav global positioning device and the same option is available for blind pedestrians.

A number of systems have been developed. One feature which they all seem to share, though, is a reliance on the use of speech, and apparent minimal use of non-speech sounds.

Some of the systems developed are:

Trekker: Based on the Maestro, which is a PDA designed to be accessible to blind users, the Trekker is a talking GPS addition. Further details are available at `http://www.humanware.ca`.

BrailleNote GPS: BrailleNote is a portable braille PDA which also has an optional GPS attachment. It displays information in speech and braille. It is also available from Humanware (`http://www.humanware.ca`), and [39] is a (somewhat dated) comparative evaluation of Trekker and BrailleNote GPS.

Sendero: This company markets accessible GPS software for a number of devices, including the BrailleNote, the VoiceSense PDA, Windows Mobile and Symbian phones as well as non-mobile devices - the desktop PC for route planning (http://www.senderogroup.com). These rely almost entirely on voice output.

Mobic: Mobic was an experimental system developed with funding from the European Union. An evaluation of the system is documented in [40], but the output was entirely spoken.

Satellite navigation is a technology developed to assist in navigation tasks by providing the user's location in the world accurate to the nearest few meters. Originally this was for military personnel operating in unfamiliar territory (presumably because they had just invaded the territory!). However, it did not take long for developers to realize that the technology might be a valuable aid for those whose navigation problems arise from their not being able to see. The systems listed above are examples of this. Further developments and improvements will no doubt occur and it would seem to be an application in which there is potential for the use of non-speech sounds. The street is an environment in which naturally occurring sounds are invaluable to a blind pedestrian. Guidance information presented in a way which is complementary to the natural soundscape would be most valuable; it must be possible to improve on plain speech.

17.6 Other Systems

There are a variety of other attempts to translate graphical materials into sounds. They are all research projects which have not yet found their way into everyday use, but brief details are given here, along with links to further information.

17.6.1 Soundgraphs

A Cartesian graph is a simple but rich visual representation of information. See Figure 17.2, for instance. Many people have had the idea that the curve of such a graph could be represented by a soundgraph or auditory graph based on a sound, the pitch of which varies with the height of the curve, but one of the first published suggestions was [41]. A number of different groups have implemented the idea, including [42, 43, 44]. Walker & Mauney present guidelines on soundgraph design in [45]. (See also chapters 2, 6, 8).

A powerful visual facility is that of the *glance*. In other words, the viewer can look briefly at a visual representation and get an overall (but imprecise) impression of its meaning. Playing the waveform of the soundgraph curve, as in Sample **S17.13**, gives a similar overall impression of the curve's shape. To gain more precise information the user might interact with the sound. Thus a sound cursor can be moved left and right along the curve and by listening for the point of highest pitch, the user might locate the maximum in the curve in Figure 17.2. Having located the point, its coordinates could be found using speech.

The soundgraph implementation of Edwards and Stevens [43] facilitated the location of such turning points (maxima and minima) by allowing the user to hear the derivative of the curve. At a turning point the slope of the curve is zero and hence its derivative is a constant. Listening to the derivative, it should have constant pitch at such a point. Grond et al. [46] have taken this idea a step further by displaying the first m terms of the Taylor Series of a

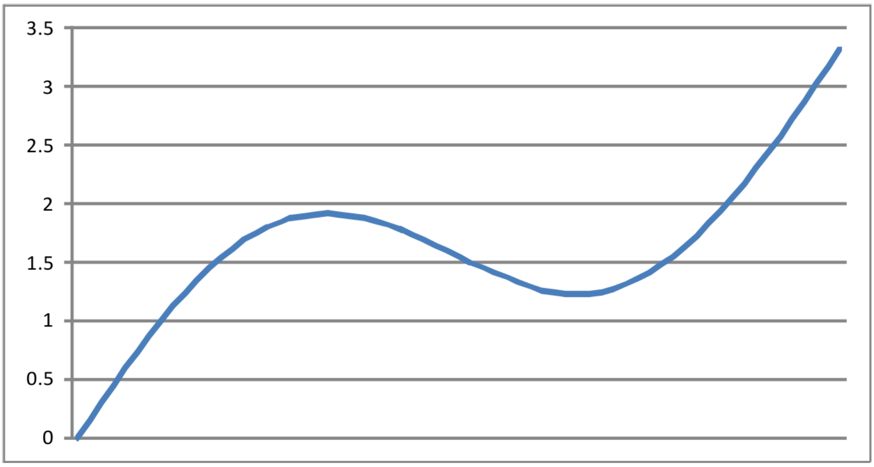

Figure 17.2: A graph of $y = \sin(x) + x/2$, a typical curve, which might be represented as a soundgraph. A sound representation of this graph can be found as Sample **S17.13**.

function. This work is in its early stages but the authors conclude that for suitable students (essentially those with experience of auditory media) 'the sonified functions are very supportive to grasp important characteristics of a mathematical function' [46, p. 20]. Examples of these sonifications are available at http://www.techfak.uni-bielefeld.de/ags/ami/publications/GDH2010-SEW.

A good soundgraph implementation should have the advantage of giving the user a feeling of very direct interaction with the *mathematics*. For instance, the user can move along the curve, sensing significant points (e.g., a maximum turning point) just by hearing the variation in the pitch (rising then falling). This contrasts with other representations such as algebra, where the requirement to manipulate the symbols can interfere with the appreciation of the mathematics that they represent. Yu, Ramloll, and Brewster [47] have gone a step further in facilitating direct interaction with soundgraphs by adding haptic interaction via the Phantom force-feedback device.

Soundgraphs have generally been used to represent curves on Cartesian graphs. Another form of graph is the scatter plot. Riedenklau et al. have developed a very novel non-visual representation of the scatter plot that uses sounds and Tangible Active Objects (TAOs) [48]. A TAO is a small plastic cube with on-board processing and wireless communication facilities. Placed on a Tangible Desk (*tDesk*) surface they can be tracked by camera. Scatter plots can be represented on the surface and the position of the TAO relative to clusters is fed back in an auditory form as a *sonogram* [49]. Sighted testers (who were blindfolded for the experiment) matched the TAO representation of different scatter plots to visual representations with a 77% success rate. Again this work is in its early stages and further developments - including testing with blind people - are proposed.

17.6.2 Audiograph

Audiograph [50, 51, 52] was a system originally designed to test how much information could be conveyed in non-speech sound, but it was soon realized that the most appropriate application would be as a means of presenting graphical information to blind people.

The following graphical information is communicated for each graphical object:

1. the position of each object;
2. the type, size and shape of each object;
3. the overall position of the objects using various scanning techniques.

All these used a similar metaphor - a coordinate point is described by using a musical mapping from distance to pitch (a higher note describing a larger coordinate value), and x and y co-ordinates are distinguished by timbre (Organ and Piano).

17.6.3 Smartsight

This is a simple form of translation from visual pixel information to non-speech sounds of different pitch, similar to soundgraphs [53, 54, 55]. An auditory cursor sweeps across the graphic horizontally. As it intersects a black pixel it makes a sound, the pitch of which represents the vertical height of the pixel. Figures 17.3 and 17.4 and their accompanying sound samples show how simple shapes are translated using this scheme.

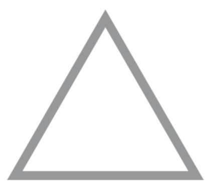

Figure 17.3: A triangle, which is rendered in sound by Smartsight as Sample **S17.14**. There is a constant (low) tone, representing the base of the triangle with rising and falling tones representing the other sides.

Simple graphics, such as those above can be perceived quite easily without training, but the developers claim that with training the same approach can be successfully used with much more complex layouts. Figure 17.5 is an example of a more complex, compound shape, but the developers claim that with training listeners can even interpret moving, animated graphics.

Smartsight originated in research in the University of Manchester, Institute of Science and Technology (now part of the University of Manchester) but was transferred to a spin-off company which has the objective of commercializing the idea. As yet, though, commercial success seems limited; during the writing of this chapter the company's web site disappeared.

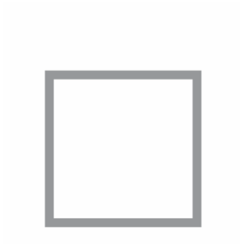

Figure 17.4: A square, which is rendered by Smartsight as Sample **S17.15**. Notice that this starts with a sharp chord, representing the left-hand vertical edge, followed by a pair of notes representing the horizontal edges and finishes with another vertical edge.

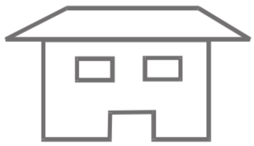

Figure 17.5: Sample graphic that would translate to Sample **S17.16** using the Smartsight system. The picture is a stylized, symmetrical house, with a trapezoid roof, two rectangular windows and a door.

17.7 Discussion

This chapter has been largely concerned with the use of hearing as a substitute for vision, but the two senses, and the stimuli which interact with them, are very different. The 'bandwidth problem' has been discussed earlier. Related to that is the fact that, in general, sight is passive. That is to say that, except in artificial conditions of darkness, visual information is always available. Thus, sighted people receive vast amounts of visual information constantly, and large parts of their brains are assigned to processing that information. Sound is also inherently temporal, while vision is more spatial. Of course sounds have a spatial origin and visual objects can move over time, but the emphasis is different in each case.

Sound, by contrast, is active, inasmuch as something must be moving to generate the sound. Most objects do not emit sounds and so to make them accessible to the auditory channel they must be made noisy. Sometimes objects can be embodied in sounds - as in screen readers which assign sounds to the elements of computer programs. Other systems work with the light analogy more directly. The Voice system operates on ambient light. It captures the visual scene, through video cameras, and converts them into sound, but in doing this it converts from the spatial domain to the temporal. In other words, the pixels are presented as an auditory raster scan, not in parallel. The problem is that the auditory sense is poorly equipped to interpret a scene thus presented.

Other devices work rather more like flashlights. The 'K' Sonar physically resembles a flashlight and it creates an auditory signal representing the portion of the environment captured in its (very limited) 'beam'. The signal presented is simple, but impoverished. With training, users can presumably interpret the sounds well (Table 17.1), but such perception is hardly comparable with vision.

The spatialized clicks of González-Mora's system are claimed to give a good picture. These are simple sounds. Serial presentation in a random order may overcome some of the problems of translation from the visual (spatial) to the auditory (spatial and temporal) domain, but the work is still experimental and yet to be proved.

17.8 Conclusion

When the sense of sight is missing other senses must be recruited. Sight is a very powerful sense and so it is very difficult (perhaps impossible) to completely substitute for it. Nevertheless, the auditory sense has great potential as an alternative. Much research effort has been expended into developing technologies which will do this, as described in this chapter. Yet it is significant that this chapter is almost solely concerned with the description of research projects; very few of the devices described are in commercial production and those which are tend to sell in small numbers (see also [56]). In other words, the great potential for the use of non-speech sounds as experienced in everyday life and highlighted by the passage from John Hull is not being realized.

This arrested development of auditory representation is a phenomenon which might be apparent in other chapters of this book; we authors and researchers know the potential for the use of auditory displays and are enthusiastic about promoting their adoption - yet the users are unconvinced. Within the context of this chapter specifically, one might expect that the users - those without sight - would in some senses be the easiest to convince, would be most willing to adopt an alternative for the sense which they lack, even if the alternative is less-than-perfect. Yet this is not the case.

Furthermore, while this chapter is concerned with disabled people, it has concentrated on those with visual disabilities. If auditory interaction has all the benefits and powers that we assert it has, then surely it could be a useful aid to those who have difficulties in interacting with technology due to other forms of impairment? Yet there seems to be almost no work that has demonstrated this to be the case.

It is not unusual for a researcher to conclude that what is required is more research, yet that is not always a cynical attempt at self-preservation. So it is in this case that there is a genuine need. We can develop auditory interfaces that are as good as the simple white cane, which give as much information as the cane (including the auditory feedback as it clicks on surfaces), and interfaces which can provide as much richness as rain falling on a garden - but as yet we do not know how.

Bibliography

[1] Hull, J., *On Sight and Insight: A Journey into the World of Blindness*. 1997, Oxford: Oneworld. 234.

[2] Stockman, T., *Listening to people, objects and interactions*, in *Proceedings of the 3rd Interactive Sonification Workshop (ISon 2010)*. 2010: Stockholm, Sweden. p. 3-8.

[3] Tiresias. *Users with disabilities - The numbers*. [Webpage] 2004 5 October 2004 [cited 2005 1 September]; Available from: http://www.tiresias.org/phoneability/telephones/user_numbers.htm.

[4] Carruthers, S., A. Humphreys, and J. Sandhu. The market for RT in Europe: A demographic study of need. in *Rehabilitation Technology: Strategies for the European Union (Proceedings of the First Tide Congress)*. 1993. Brussels: IOS Press.

[5] Elkind, J., The incidence of disabilities in the United States. *Human Factors*, 1990. 32(4): p. 397-405.

[6] Bruce, I., A. McKennell, and E. Walker, *Blind and Partially Sighted Adults in Britain: The RNIB Survey*. 1991, London: HMSO.

[7] Weber, G., Reading and pointing - New interaction methods for braille displays, in *Extra-ordinary Human-Computer Interaction: Interfaces for Users with Disabilities*, A.D.N. Edwards, Editor. 1995, Cambridge University Press: New York. p. 183-200.

[8] Weber, G., Braille displays. *Information Technology and Disability*, 1994. 1(4).

[9] Blenkhorn, P., Requirements for screen access software using synthetic speech. *Journal of Microcomputer Applications*, 1993. 16(3): p. 243-248.

[10] Edwards, A.D.N., *Adapting user interfaces for visually disabled users*. 1987, Open University.

[11] Edwards, A.D.N., Soundtrack: An auditory interface for blind users. *Human-Computer Interaction*, 1989. 4(1): p. 45-66.

[12] Mynatt, E.D. and G. Weber. Nonvisual presentation of graphical user interfaces: Contrasting two approaches. in *Celebrating Interdependence: Proceedings of Chi '94*. 1994. Boston: New York: ACM Press.

[13] Gaver, W., Auditory Icons: Using sound in computer interfaces. *Human Computer Interaction*, 1986. 2(2): p. 167-177.

[14] Downie, A. *VoiceOver screen reader for Apple Macintosh Leopard*. 2008 April 2008 [cited 2011 2 February]; Available from: http://accessiblecli.wordpress.com/2009/07/28/voiceover-screen-reader-for-apple-macintosh-leopard/.

[15] Leventhal, J., Not what the doctor ordered: A review of Apple's VoiceOver screen reader. *AccessWorld*, 2005. 6(5).

[16] Hull, J., *Touching the Rock: An Experience of Blindness*. 1990, London: SPCK. 165.

[17] McGrath, R., T. Waldmann, and M. Fernström. Listening to rooms and objects. in *Proceedings of AES 16th International conference on Spatial Sound Reproduction*. 1999. Rovaniemi, Finland.

[18] Bradley, N.A. and M.D. Dunlop. Investigating context-aware clues to assist navigation for visually impaired people. in *Proceedings of Workshop on Building Bridges: Interdisciplinary Context-Sensitive Computing*. 2002. Glasgow: University of Glasgow.

[19] Walker, B.N. and J. Lindsay. Navigation performance in a virtual environment with bonephones. in *Proceedings of the 11th International Conference on Auditory Display (ICAD 2005)*. 2005. Limerick, Ireland.

[20] Massof, R.W. Auditory assistive devices for the blind. in *Proceedings of ICAD 2003 (International Conference on Auditory Display)*. 2003. Boston.

[21] Yen, D.H. *Currently Available Electronic Travel Aids for the Blind*. 2005 [cited 2006 3 April]; Available from: http://www.noogenesis.com/eta/current.html.

[22] Brabyn, J.A., New developments in mobility and orientation aids for the blind. *IEEE Trans. Biomedical Engineering*, 1982. BME-29(4): p. 285-289.

[23] Shoval, S., I. Ulrich, and J. Borenstein, Computerized obstacle avoidance systems for the blind and visually impaired, in *Intelligent Systems and Technologies in Rehabilitation Engineering*, H.-N.L. Teodorescu and L.C. Jain, Editors. 2001, CRC Press: Boca Raton. p. 413-447.

[24] Borenstein, J. The NavBelt - A Computerized Multi-Sensor Travel Aid for Active Guidance of the Blind. in *Proceedings of the Fifth Annual CSUN Conference on Technology and Persons With Disabilities*. 1990.

[25] Shoval, S., J. Borenstein, and Y. Koren. The Navbelt - A Computerized Travel Aid for the Blind. in *Proceedings of the RESNA '93 Conference*. 1993.

[26] Shoval, S., J. Borenstein, and Y. Koren. Mobile Robot Obstacle Avoidance in a Computerized Travel Aid for the Blind. in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. 1994. San Diego.

[27] Shoval, S., J. Borenstein, and Y. Koren, Auditory Guidance With the NavBelt - A Computerized Travel Aid for the Blind. *IEEE Transactions on Systems, Man, and Cybernetics*, 1998. 28(3): p. 459-467.

[28] Kay, L., A sonar aid to enhance spatial perception of the blind: Engineering design and evaluation. *Radio and Electronic Engineer*, 1974. 44(11): p. 605-627.

[29] Meijer, P.B.L., An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 1992. 39(2): p. 112-121.

[30] Trivedi, B., Sensory hijack: rewiring brains to see with sound. *New Scientist*, 2010 (2773).

[31] González-Mora, J.L., et al., Seeing the world by hearing: Virtual acoustic space (VAS) a new space perception system for blind people, in *Touch Blindness and Neuroscience*, S. Ballesteros and M.A. Heller, Editors. 2004, UNED: Madrid, Spain. p. 371-383.

[32] Miller, L., Diderot reconsidered: Visual impairment and auditory compensation. *Journal of Visual Impairment and Blindness*, 1992. 86: p. 206-210.

[33] Rauschecker, J.P., Compensatory plasticity and sensory substitution in the cerebral cortex. *TINS*, 1995. 18(1): p. 36-43.

[34] Sadato, N., et al., Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, 1996. 380(6574): p. 526-528.

[35] Veraart, C., et al., Glucose utilisation in visual cortex is abnormally elevated in blindness of early onset but decreased in blindness of late onset. *Brain Research*, 1990. 510: p. 115-121.

[36] Zangaladze, A., et al., Involvement of visual cortex in tactile discrimination of orientation. *Nature*, 1999. 401: p. 587-590.

[37] Gonzalez-Mora, J.L., et al. Seeing the world by hearing: Virtual Acoustic Space (VAS) a new space perception system for blind people. in *Information and Communication Technologies, 2006. ICTTA '06. 2nd*. 2006.

[38] Tyler, R.S., ed. *Cochlear Implants: Audiological Foundations*. 1993, Singular: San Diego. 399.

[39] Denham, J., J. Leventhal, and H. McComas. *Getting from Point A to Point B: A Review of Two GPS Systems*. 2004 [cited 2006 24 July]; Volume 5 Number 6:[Available from: http://www.afb.org/afbpress/pub.asp?DocID=aw050605].

[40] Strothotte, T., et al. Development of dialogue systems for a mobility aid for blind people: Initial design and usability testing. in *Assets '96*. 1996. Vancouver: ACM.

[41] Mansur, D.L., M. Blattner, and K. Joy, Sound-Graphs: A numerical data analysis method for the blind. *Journal of Medical Systems*, 1985. 9: p. 163-174.

[42] Brown, L. and S. Brewster. Design guidelines for audio presentation of graphs and tables. in *Proceedings of the 9th International Conference on Auditory Display (ICAD 2003)*. 2003. Boston.

[43] Edwards, A.D.N. and R.D. Stevens, Mathematical representations: Graphs, curves and formulas, in *Non-Visual Human-Computer Interactions: Prospects for the visually handicapped*, D. Burger and J.-C. Sperandio, Editors. 1993, John Libbey Eurotext: Paris. p. 181-194.

[44] Sahyun, S.C. and J.A. Gardner. Audio and haptic access to Maths and Science: Audio Graphs, Triangle, the MathPlus Toolbox and the Tiger Printer. in *Computers and Assistive Technology, ICCHP '98: Proceedings of the XV IFIP World Computer Congress*. 1998. Vienna & Budapest.

[45] Walker, B.N. and L.M. Mauney, Universal Design of Auditory Graphs: A Comparison of Sonification Mappings for Visually Impaired and Sighted Listeners. *ACM Trans. Access. Comput.* 2(3): p. 1-16.

[46] Grond, F., T. Droßard, and T. Hermann, *SonicFunction: Experiments with a function browser for the visually impaired*, in *Proceedings of the The 16th International Conference on Auditory Display (ICAD 2010)*. 2010. p. 15-21.

[47] Yu, W., R. Ramloll, and S.A. Brewster, Haptic graphs for blind computer users, in *Haptic Human-Computer Interaction*, S.A. Brewster and R. Murray-Smith, Editors. 2001, Springer-Verlag: Heidelberg. p. 102-107.

[48] Riedenklau, E., T. Hermann, and H. Ritter, *Tangible active objects and interactive sonification as a scatter plot alternative for the visually impaired*, in *Proceedings of the The 16th International Conference on Auditory Display (ICAD 2010)*. 2010. p. 1-7.

[49] Hermann, T. and H. Ritter, Listen to your data: Model-based sonification for data analysis, in *Advances in intelligent computing and multimedia systems*, G.E. Lasker, Editor. 1999, International Institute for Advanced Studies in System Research and Cybernetics. p. 189-194.

[50] Rigas, D., *Guidelines for auditory interface design: An empirical investigation*, in *Department of Computer Science*. 1996, Loughborough University: Loughborough.

[51] Rigas, D.I. and J.L. Alty. The use of music in a graphical interface for the visually impaired. in *Proceedings of Interact '97, the International Conference on Human-Computer Interaction*. 1997. Sydney: Chapman and Hall.

[52] Alty, J.L. and D.I. Rigas. Communicating graphical information to blind users using music: The role of context. in *Making the Impossible Possible: Proceedings of CHI '98*. 1998: ACM Press.

[53] Cronly-Dillon, J.R., et al., Blind subjects analyse photo images of urban scenes encoded in musical form. *Investigative Ophthalmology & Visual Science*, 2000. 41.

[54] Cronly-Dillon, J.R., K.C. Persaud, and R.P.F. Gregory, The perception of visual images encoded in musical form: a study in cross-modality information transfer. *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, 1999. 266: p. 2427-2433.

[55] Cronly-Dillon, J.R., K.C. Persaud, and R. Blore, Blind subjects construct conscious mental images of visual scenes encoded in musical form. *Proceedings of The Royal Society of London Series B-Biological Sciences*, 2000. 267: p. 2231-2238.

[56] Thierry, P., et al., Image and video processing for visually handicapped people. *J. Image Video Process.*, 2007. 2007(5): p. 1-12.