SONIFYING THE LONDON UNDERGROUND REAL-TIME-DISRUPTION MAP

Louise Valgerður Nickerson, Tony Stockman and Jean-Baptiste Thiebaut

Interaction Media and Communication Group
Department of Computer Science
Queen Mary University of London
Mile End Road, London E1 4NS
United Kingdom
{lou,tonys,jbt}@dcs.qmul.ac.uk

ABSTRACT

In mobile computing, there is a need for interfaces that better suit the context of use. Auditory interfaces have the potential to address the limitations of small screens and support eyes-free tasks. In order to fill this gap, we must develop more fluid and usable auditory interfaces. A key aspect of this is understanding the process of designing overviews. In this work, we describe a conceptual strategy for providing an overview of disruptions in the London Underground: The approach adopted is based on what information is perceived as most crucial to the user.

[Keywords: Auditory interfaces, overviews, sonification, accessibility, mobile computing.]

1. INTRODUCTION

Imagine this: you have an important meeting at work first thing in the morning. Running a few minutes late, you rush out the door, coffee in hand. Halfway to the bus stop, you realise you should have checked your local transport links so you know whether your regular route is problem-free. Switching your coffee to your other hand, you pull out your mobile phone and browse to the transport information site to check for problems reported in the system. Slowing down to type in the web address, you miss the light for crossing the street; as you follow links you glance up occasionally to avoid other pedestrians. As you scroll the online map on your tiny screen, you decide it might be quicker to call your husband who is still at home and ask him to look it up for you instead. He reports that your local metro station is closed so you'd best take the bus to the next one but luckily there are no problems reported otherwise. You continue your way to work...

The above scenario highlights many problems: entering text while walking, getting information in a timely manner on a mobile device, efficient navigation of web content, etc. There are many occasions when mobile computing could prove beneficial, for example, location-based services. However, the interaction problems indicate a different approach is necessary to make them more usable and better aligned to the context of use.

What makes creating efficient and effective interactions difficult is the difference between desktop computers and mobile ones, notably the lack of screen space, the lower processing power, slower network connections, and the different input devices (5-way navigation, stylus, thumb keyboard, etc.). Another factor that needs to be taken into consideration is the state of the user, the type of device and the environment in which the interface is used. In particular, users are often impatient, distracted, and/or performing other tasks and require a fast, simple interface that does not distract them from other things they may be doing at the time.

In order to begin to address some of the issues mentioned above, we investigate the use of sound to address the limitations of small screen size on mobile devices. Specifically, this work explores how sound can help improve how information is presented. There have been many attempts to improve the presentation of information on small screens via advanced visualisation techniques [1]. An alternate solution is to fully or partially offload the presentation to another modality. Sound is particularly attractive because it is independent of device size. It is also eyes-free, allowing the user to carry on with their current task. In the case of mobile phones and portable music players, audio output is an integral part of the device and therefore easy to take advantage of.

However, in order to effectively present information in the auditory realm, current interfaces need to mature. Typical auditory interfaces use a lot of speech and tend to be serial in presentation. Screen-readers, for example, read in order from left to right and top to bottom through the information on the screen. This often results in slow interaction and exploration can be tedious. Looking to Shneiderman's info seeking principle [2], later adapted for auditory interaction [3], both interaction and overviews are very important for exploration of interfaces.

We present the design of an auditory version of the London Underground real-time disruption map [4], optimised for a mobile phone. The Underground, or Tube, is London's metro system and the real-time disruption map (disruption map) is part of the real-time information that the Tube's governing body, Transport for London (TfL) publishes about the current state of the public transportation network in London. The disruption map presents information that is useful while mobile, but in its current state it is very hard to use on a phone. Our proposed interface provides interaction in the form of a three-level hierarchy of information. The first level of this hierarchy is the overview described in this paper. The overview design presents the whole of the system and its interconnections using concurrent non-speech sound, drawing on what information is perceived to be the most important. We discuss the foundations for the design and in particular discuss its advantages and drawbacks against several usability criteria.

1.1. Background

One reason for the interest in auditory interfaces is to provide access to graphical user interfaces to the visually-impaired. Screen-readers such as JAWS for Windows [5] and Window-Eyes [6] are the main accessibility tools used by the visually impaired. They are speech-based and read linearly through the items in the interface. Even though what is read can be tailored by changing verbosity settings, this interaction is often slow for real world task requirements and inflexible, there has long been interest in providing a better interaction model.

1.1.1. Non-speech presentation strategies

Auditory icons [7] and earcons [8] are the main non-speech strategies for presenting information in non-speech sound. Auditory icons are short real world sounds that are designed to be mapped to actions or events. Earcons are short abstract sounds that can be encoded with information based on their pattern and timbre. The advantage of auditory icons is that, if perceived as a signifier and linked to the signified, they are easily remembered. However related to the real world sound, they rely on a strong metaphor and there is not always a logical mapping for all that one might want to represent. Earcons, being abstract, have a higher learning curve but the potential to be powerful in describing relationships between objects. Another more recent presentation strategy is spearcons [9] which are based on speeding up speech to the point of unintelligibility in order to create abstract sounds.

Earcons have been used in such work as providing feedback in telephone hierarchies [10] in which earcons were shown to aid in determining location in the menu hierarchy. An example of a relatively sophisticated design for providing improved auditory interaction was the Mercator Project [11] for X Windows, which proposed a combination of speech and auditory icons to map visual objects to audio. Spearcons have been used in auditory menus to aid in their ordering and faster access. In the sonification of UML diagrams [12], simple abstract sounds describe the relation between various elements in UML diagrams and provide information about their content.

1.1.2. Auditory overviews

Little work has been done on auditory overviews and this area presents a gap in auditory interface research. The necessity thereof is highlighted by the Auditory Information Seeking Principle [3] (overview, navigate, filter then details-on-demand). The principle is applied in the iSonic project [13] which provided an auditory interface to US census data. Users navigate a map of the United States and tons are used to indicate the population of the state. The interface uses a combination of speech and non-speech, where the speech was used for details and non-speech for overview information. Another project, AudioView [14], looks at providing overviews of Java source code using speech and earcons. The overview is provided through earcons which represent the types of control structures in the source code.

Both the above projects use serial (or sequential information) presentation, that is, no sounds overlapped. In more complex interfaces where there are interconnections to be displayed, concurrent presentation is attractive. Concurrent presentations have the advantage of speed, but can suffer from loss of coherency. McGookin and Brewster [15] suggest that concurrent earcons that have well-separated timbres and that are staggered by 300ms have a better chance of being identified. In order to pursue the possibilities of concurrency, we also draw upon auditory scene analysis [16] where sounds that were sufficiently separated in frequency and sounded sufficiently different were perceived to be different streams.

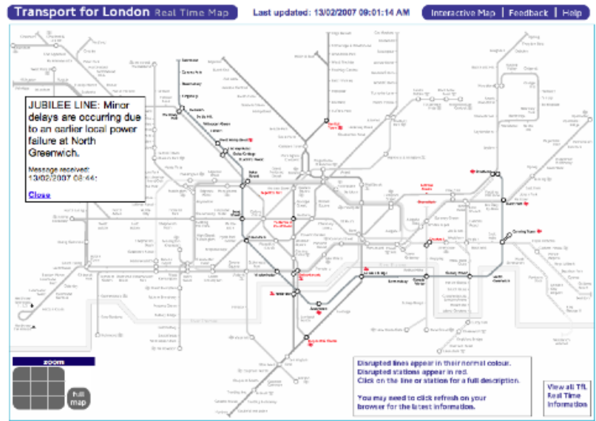

Figure 1: *In this view of the disruption map there are a number of disrupted stations (in red) and one disrupted line (Jubilee line running north-west to east). The disruption message reads: JUBILEE LINE: Minor delays are occurring due to an earlier power failure at North Greenwich.*

2. THE REAL-TIME DISRUPTION MAP

The disruption map is based on the standard Tube map and displays how the service is running. The standard map is a simplified and stylised map of the Tube network where each line is associated with a colour. On the disruption map, lines that are working properly are greyed out, leaving only the disrupted lines in colour. Stations on disrupted lines are displayed in black, but if the station itself is disrupted, it appears in red. Users can see immediately how well the Tube network is running. The map also provides access to details of the disruptions. Clicking on a disrupted line (shown in colour) or station (shown in red) pops up a window with information about the problem and time it was recorded. An example of a disruption message is shown in Figure 1. However, on a mobile device, the image is either shrunk, does not display or requires a lot of scrolling. The disruption messages appear as a separate list on the page. The result is a less usable and intuitive interface.

2.1. Breakdown

The key parts of the disrupted portions of the Tube system and the disruption messages. The key functions of the map are the immediate view of state of entire network and access to the disruption messages. It is also key to represent the relationship between the lines. TfL reports disruptions for stations, lines and partial lines. We define the system as comprising paths (Tube lines) that have sections (ranging from a single station to a series of stations). A state of disrupted or normal can be applied to the entire line or a subsection. Stations can belong to one or more lines and can equally have a state, of which there are three. A station can be disrupted on its own or as part of a disrupted line, the third state being normal. The disruption messages are made up of a description and a time-stamp.

level	overview	line	message
description	the state of the system displaying lines with problems	the state of a single line	the disruption message
press up navigates to	nothing	overview	line
press down navigates to	line level	disruption message	nothing
press left navigates to	nothing	previous disrupted line	previous disruption message
press right navigates to	nothing	next disrupted line	next disruption message
keypad entry navigates to	line requested	line requested	line requested

Table 1: *Interaction in the conceptual design and the actions required to move from level to level.*

2.2. Alternatives to the disruption map

2.2.1. Text-based real-time news

Users can alternatively use a textual description of the problems occurring on the Tube. In the textual description, information is presented sequentially. The lines with problems in a section of the line or at multiple stations are listed. Below this list, there is a second one displaying all problems per station. This means that there is duplication of information on the page: a problem would be listed once per line and a second time per station.

2.2.2. Telephone hotline

TfL also provides an audio service that speaks the disruptions on the underground. A call made on a Sunday evening took 4'30 for all the information to be presented. This audio system is not interactive and information is presented sequentially.

2.3. An example usage of the disruption map

We determine that a user will approach the map with the question *am I affected?* This is further broken down into *is a line on my journey affected?* and *is a start, end or transfer station on my journey affected?*. If the answer to any of these questions is yes, the question becomes *how much am I affected?* Below is an example usage of the map with a novice user asking those questions.

You want to go from a point A to a point C, transferring at point B. You are thus interested in problems on each metro line and key station on your journey. You connect to the TfL disrupted map. The Tube map is displayed on the screen, and you already (vaguely) know the locations of the stations in question. The first thing you want to locate is the line you take first. As you look for it, you notice that the line is grey and pale instead of coloured, while some other lines are coloured. At your start station (point A), you use the mouse to check the interactive options. As nothing changes when you put the mouse on the station location, nothing seems to be available. Then, following the metro line to point B where you change and then along the second line, also greyed out, to point C which is your destination. There, the station name is coloured. You click on it, and the map displays a message indicating a disruption - reduced escalator service - which does not affect your journey. You quit the service after approximately 40 seconds.

From this use case, we observe that even novice users of the disruption map may access the desired information fairly quickly and do not encounter a high learning curve. The map itself is already familiar so the location of the key points is easy. This constitutes the first task in this user case. The second task was to reach virtually the transfer station, and then go to the end of the journey. At the beginning of this task, you are focused on the portion of line you will be travelling along. A more experienced user of the map would have faster interaction speeds due to knowledge of the interaction; they would be able to skip finding their line by only paying attention to highlighted information (information in colour).

3. AN AUDITORY CONCEPTUAL DESIGN OF THE DISRUPTION MAP

The conceptual design presented here aims to speed up auditory access to information by tackling two missing aspects of the interface: overview and interaction. We shall outline the design of both of these components of the interface, before going on to describe in detail the development and evaluation of the overview component. Going back to the Auditory Information Seeking Principle [3]: overview, navigate, filter then details-on-demand, we adopt a hierarchical structure that can be navigated. The interface consists of three levels: overview of all the disruptions in the system at the top, individual disruption information per line in the middle and detailed disruption messages at the bottom. As such, the second level of the interface provides navigation and filtering while the third level satisfies details-on-demand.

The hierarchy is intended to provide the information in more or less the same way that the visual domain does. That is, provide global information first (the state of the system) and navigate to details if needed (the state of a station or line). Upon initialisation of the application, the user is presented with the overview of the system. The main method of interaction would be the 5-way navigation, using up and down to move from one level to the next and left and right to navigate the current level. Additionally, we provide shortcuts to each line by way of the numeric keypad: typing the 2 first letters of the name of a line would jump to the middle level of the hierarchy for the line in question. I.e., typing 66 (or 'no') accesses the Northern Line. The interaction is further described in table 1. At the line level, navigating left and right is circular so that upon reaching the last disrupted line, pressing right would bring the user back to the first disrupted line. At the disruption level, the user can only navigate left and right through disruptions that belong the the line selected at the line level Therefore, in order to hear disruption messages for another line, the user must go back up and go to the desired line and descend into the message level. The following sections discuss in detail the design of the auditory overview of the disruption map.

3.1. Design of the overview

To provide an overview, we display concurrent information about the system. The goal is to link the overview to the schematic layout of the system. One important aspect of the Tube network is the interconnection between the lines. Therefore, each line is sonified

Figure 2: *The East London Line is defined by its start and end stations (Whitechapel, New Cross Gate and New Cross) as well as its transfer stations (Shadwell and Canada Water). As it branches, it is also defined by where the two branches meet (Surrey Quays)*

individually but the stations that interconnect with other lines coincide. In parallel with the disruption map itself, this overview only plays lines that are disrupted. If no line is disrupted, a low comfort noise/tone is played simply to reassure the user that the program has not unexpectedly quit. If only a section or station on a line is disrupted, only that portion is presented. By doing so, we convey to the user more information about the scope of the problem.

3.1.1. Scheduling of the lines

In order to properly display all the Tube lines, timings were worked out very carefully to ensure that if all lines were disrupted there would be sufficient spacing between the individual sound events. We chose to represent each line by its start and end stations and all transfer stations along the line. Figure 2 provides an example of how this was done for the East London line. At each defining station, a event sounds creating a similar but unique pattern for each line.

Initial timings were determined for each line by using the distance between the defining stations and lines were for the most part deemed to begin in the west and travel east, or begin in the north and travel south. After determining the timings, each line was merged one-by-one into a single timeline. The timings are driven by the Circle line as it shares its entire track with other lines. Additionally, all but one line intersect with the Circle line at some point, creating the additional dependancies in the timings. Essentially the whole map was flattened and stretched in places to ensure all stations occurred in the correct order.

3.1.2. The resulting overview

In order to stagger the sound events that define a completely disrupted Tube system, we turn to McGookin and Brewster [15] and Bregman [16]. McGookin and Brewster recommend staggering of concurrent earcons by 300ms to ensure intelligibility. However as the events are quite short and do not have a pattern to be discerned, 20-50ms was deemed sufficient. This is upheld by concepts from Bregman that show that similar sounds are grouped into streams. The resulting stream or pattern of each line is what needs to stand out rather than the individual events. The next step was to make each Tube line sound different. Each line was logarithmically spaced in a frequency space ranging from 220Hz to 1500Hz. The frequency separation allows the user to group sounds as belonging together. Lines that branch deviate slightly from their frequency to indicate the two paths the line is following. Identifying a line by its frequency is further reinforced by assigning each line a timbre. The resulting overview therefore defines a line by its pattern (defining stations), sound (frequency and timbre) and when it starts and ends (timing). Figure 3 shows the start and end timings for each line. With up to 6 lines disrupted the overview is relatively comprehensible however beyond that, it becomes cacophonous. At that point, the user mostly understands that there are many problems on the Tube, which is of itself useful information for a traveller.

3.1.3. Overview repetition at the line and disruption message level

At the line level, the overview sound that plays behind the speech is the overview of the line without any of the other disrupted lines playing. Comfort noise is played for any portion of the line that has no disruptions. In other words, if only stations are disrupted, they play as events but are played in the correct place in their timeline so that users can more easily recognise the location of the disruption. At the disruption message level, only the disruption being presented plays.

3.2. Line and disruption levels

As the overview is the main focus of the work, the line and disruption message levels of the interface have been kept simple. At the line level, the name of the line affected is spoken with the overview representation of the line played in the background. The repetition of the overview of the line is intended to aid the user in learning the sounds. At the disruption message level, the message is spoken in a short hand manner, presenting first the most important information: the scope of the problem (section of the line or station name) and the nature of the problem (e.g. 'delays due to signal failure at Monument' or 'reduced escalator service') followed by any more minor details (e.g. 'until April 2007' in the case of an escalator disruption). Once again, the sound of the disruption from the overview is played to help the user make the right association.

4. IMPLEMENTATION AND INITIAL EVALUATION

4.1. Initial prototype

The overview described in previous sections has been built into an initial prototype. The sounds were created using SuperCollider3[1] and SimpleSynth[2] to produce MIDI patterns for each line. A test-

[1] http://supercollider.sourceforge.net/
[2] http://pete.yandell.com/software/

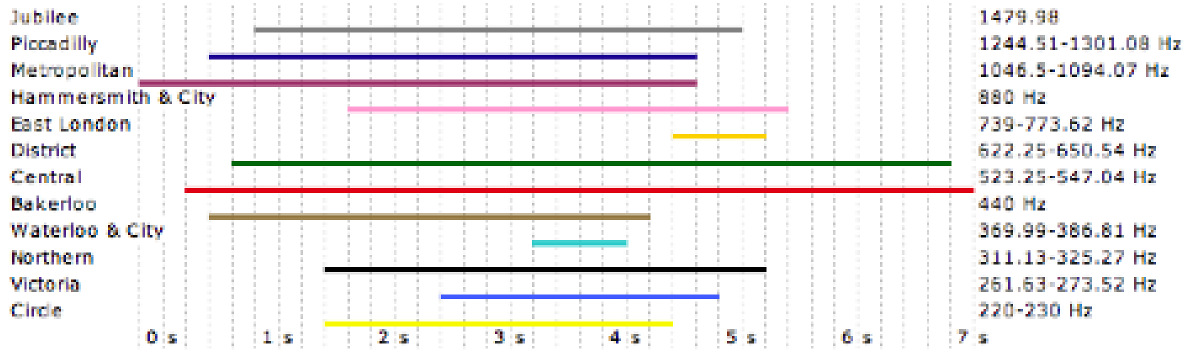

Figure 3: *The timings and frequencies for the Overview*

ing interface has been developed in SVG[3] allowing the system to be demonstrated alongside the visual representation. A demo can be seen and heard on the aIMC project page.

4.2. Initial and future evaluation

We aim to evaluate the resulting overview on several levels:

clarity of the presentation We define clarity as the level of ambiguity a user feels about the information presented.

ease of use Ease of use refers to the straight-forwardness of the interface.

suitability to the mobile context The overview needs to be tolerant of ambient noise and distracted users.

speed The presentation of the overview is always approximately 7 seconds which seems a sufficiently short period of time given the level of information being presented.

predictability We define predictability as the extent to which the interface behaves consistently. The user will be able to rely on the interface to take the same amount of time each time but for the number of streams to change with time.

learning curve It is hypothesised that the approach to overview presentation will require something of a learning curve in order for users to become familiar with the presented sounds.

support for advanced users A good interface needs to support not only novice users but support increased familiarity with the interface. However, this investment in learning time could be rewarded due to the level of information encoded in the overview, as there is the potential for grasping the scope of the disruption without having to explore further. We anticipate that the approach taken here to overview presentation will provide a good level of support for advanced users.

4.2.1. Informal interviews

To date, we have conducted several informal interviews to discover ways the overview could be refined. Using the testing interface, we have asked three users to tell us their impression of the interfaces. The aim was mainly to address the first of our criteria: *clarity of the presentation*.

[3]'urlwww.w3.org/Graphics/SVG

The users reported that the different frequencies were helpful in distinguishing one Tube line from another. One in particular commented that he had not thought that frequency would be enough to separate out streams of information but was pleasantly surprised. However, one user who had the opportunity to compare an early prototype using simple tones to the current prototype with different instruments assigned to each line found it harder to distinguish the lines. Based on this feedback, we are looking at assigning timbres that are more orthogonal in order to minimise interference with each other.

Another important comment was the busyness of the overview. Currently, if a line is disrupted, all defining stations (as described in section 3.1.1) are played and each line can range from 2 to 25 events. We are simplifying the overview by reducing the number of events that define a line and are currently conducting tests where transfer stations are ignored unless they intersect with one or more other disrupted lines. Fewer events also gives us the possibility of reducing the length of the overview.

5. DISCUSSION

The overview described in this work has the goal of presenting not only the information but also the structure of what is being presented. The user will get an idea of how the Tube lines relate to each other represented entirely in non-speech audio. The overview provides a concurrent presentation of the information that is perceived to be most important. As discussed in section 2.1 the key information is the disruptions and their interconnections. By analysing the Tube network, it is apparent that disruptions can cascade due to the relationships between the different lines. If we de-couple the disruptions and lines from the underlying structure, the side effects of disruptions become less obvious and it is left to the user to use her knowledge of the network to grasp the implications of a particular disruption. As such, we chose to expose the relationships rather that deem them of lesser importance.

5.1. Future work

The next step in this design is the formal evaluation of the approach taken to overview presentation against the criteria above. However, it is already possible to think of next steps we may take. One such step is attempting to display the severity of the disruption in the overview. This could be done to help the user get a measure of how much their journey could be affected prior to exploring the disruption messages. A simple method of doing this would be to make more severe problems louder. Perceptually, a louder sound is perceived as closer, thus expressing the urgency of the disruption. Care would need to be taken in order to ensure such an increase in amplitude does not mask minor disruptions. Currently severity of disruptions are not displayed in the visual disruption map but are in the text-based enumeration of disruptions.

Along with the improvements being made on the basis of the initial evaluation, another issue worth attention in our overview design are how to emphasise station disruptions so that they do not get lost in the presentation of other, longer disruptions . For the emphasis of station disruptions, amplitude could be used but is perhaps better suited metaphorically to expressing severity. Another method of drawing attention is to make the sound event longer and have the frequency jitter in order to draw attention to it. Experimentation is required to determine if this is a viable approach.

6. CONCLUSION

This paper discusses an approach to creating an overview for the Transport for London real-time disruption map. The interface is designed for use on a mobile phone and the interaction is based on the input devices available: numeric keypad and 5-way navigation. Mobile computing is a context in which auditory interfaces could address some of the current challenges, namely the lack of screen space and the support of multi-tasking. However, before auditory interfaces can be adopted, their weaknesses must also be addressed. To this end, we have designed a conceptual interface that follows the concepts of the Auditory Information Seeking Principle [3]. The hierarchy provides filtering and easy navigation, leaving the user free to decide where to explore without having to listen to long auditory presentations. The main focus, however, is how to generate an appropriate overview. This is a subject that has not seen a lot of research. We have evaluated the approach taken in the design of the overview in section 5 and have hypothesised about its strengths and weaknesses. Future work will involve a more detailed implementation of the design, and a formal evaluation will determine if our suppositions hold.

7. REFERENCES

[1] Sami Baffoun and Jean-Marc Robert, "Etat de l'art des techniques de présentation d'information sur écran d'assistant numérique personnel," in *IHM '06: Proceedings of the 18th international conference on Association Francophone d'Interaction Homme-Machine*, New York, NY, USA, 2006, pp. 27–34, ACM Press.

[2] Ben Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proc. Visual Languages*, September 1996.

[3] Haixia Zhao, Catherine Plaisant, Ben Shneiderman, and Ramani Duraiswami, "Sonification of geo-referenced data for auditory information seeking: Design principle and pilot study," in *Proceedings of the 10th International Conference on Auditory Display*, Sydney, Australia, July 6-9 2004.

[4] Transport for London, "The real time disruption information map," http://journeyplanner.tfl.gov.uk/im/RD-T.html.

[5] Freedom Scientific, "Jaws for windows," http://www.freedomscientific.com/fs_products/-JAWS_HQ.asp.

[6] GW Micro, "Window-eyes," http://www.gwmicro.com/-Window-Eyes/.

[7] William W. Gaver, *Everyday listening and auditory icons*, Ph.D. thesis, University of California, San Diego, 1988.

[8] Meera M. Blattner, Denise A. Sumikawa, and Robert M. Greenberg, "Earcons and icons: their structure and common design principles," in *Human Computer Interaction*, 1989, vol. 4, pp. 11–44.

[9] Bruce N. Walker, Amanda Nance, and Jeffrey Lindsay, "Spearcons: speech-based earcons improve navigation performance in auditory menus," in *Proceedings of the 12th International Conference on Auditory Display*, Tony Stockman, Louise Valgerður Nickerson, Christopher Frauenberger, Alistair D. N. Edwards, and Derek Brock, Eds., June 2006.

[10] Stephen A. Brewster, "Using nonspeech sounds to provide navigation cues," *ACM Trans. Comput.-Hum. Interact.*, vol. 5, no. 3, pp. 224–259, 1998.

[11] Elizabeth D. Mynatt, *Transforming graphical interfaces into auditory interfaces*, Ph.D. thesis, Georgia Institute of Technology, April 1995.

[12] Oussama Metatla, Nick Bryan-Kinns, and Tony Stockman, "A model for structuring UML class diagrams to support non-visual interpretation and navigation," in *Proceedings of The 20th BCS HCI Group conference*, Bob Fields, Tony Stockman, Louise Valgerður Nickerson, and Patrick G. T. Healey, Eds., London UK, September 2006, vol. 2.

[13] Haixia Zhao, Ben Shneiderman, and Catherine Plaisant, "iSonic: Interactive data sonification for blind users," http://www.cs.umd.edu/hcil/audiomap/, 2006.

[14] J. Louise Finlayson and Chris Mellish, "The "audioview" – providing a glance at java source code," in *Proceedings of the 11th International Conference on Auditory Display*, Limerick, Ireland, July 6-9 2005.

[15] David K McGookin and Stephen A Brewster, "An investigation into the identification of concurrently presented earcons," in *Proceedings of the 9th International Conference on Auditory Display*, Boston, Massachussetts, United States, 6-9 July 2003.

[16] Albert S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, The MIT Press, September 1994.