Overviews and their effect on interaction in the auditory interface.

Nickerson, Louise Valgerður

For additional information about this publication click this link.
http://qmro.qmul.ac.uk/jspui/handle/123456789/8687

Overviews and their effect on interaction in the auditory interface

Louise Valgerður Nickerson

Submitted for the degree of Doctor of Philosophy

Queen Mary, University of London

2013

Declaration

I hereby declare that this thesis has not been submitted, either in the same or different form, to this or any other university for a degree. Additionally, the work submitted in this dissertation is the result of my own investigation, except where otherwise stated.

Signature:

Louise Valgerður Nickerson

Overviews and their effect on interaction in the auditory interface

Louise Valgerður Nickerson

Abstract

Auditory overviews have the potential to improve the quality of auditory interfaces. However, in order to apply overviews well, we must understand them. Specifically, what are they and what is their impact? This thesis presents six characteristics that overviews should have. They should be a structured representation of the detailed information, define the scope of the material, guide the user, show context and patterns in the data, encourage exploration of the detail and represent the current state of the data. These characteristics are guided by a systematic review of visual overview research, analysis of established visual overviews and evaluation of how these characteristics fit current auditory overviews.

The second half of the thesis evaluates how the addition of an overview impacts user interaction. While the overviews do not improve performance, they do change the navigation patterns from one of data exploration and discovery to guided and directed information seeking. With these two contributions, we gain a better understanding of how overviews work in an auditory interface and how they might be exploited more effectively.

Contents

List of Figures

List of Tables

Acknowledgements

I am very grateful for all the support I have received throughout my thesis.

Extra special thanks go to my family for being so patient and most of all to Bob Rosenberg, who was a constant source of inspiration, an excellent sounding board and super efficient proof-reader. In this category are also Tony Stockman and Nick Bryan-Kinns for all their support over the years.

I would also like to thank the IMC research group, C4DM research group and the ICAD community. Special thanks go to Greg Mills for interminable discussions on statistics and the direction of my research; Chris Frauenberger, for stretching my thought processes when I needed it most; George Papatzanis, for dragging me away from research and back to it when necessary; Chrystie Myketiak, Steve Welburn, Paul Hodkinson and Brock Craft, for being extremely thorough reviewers; Dan Stowell, for providing statistics, SuperCollider and sound design help; Rosemary Bailey for helping me understand the intricacies of statistical significance; Thomas Hermann, for showing me how exciting research can really be; and Peter Smith for helping me through administrative hurdles in the Research Degrees Office.

Chapter 1

Introduction

There are many reasons why one might augment an interface with sound or produce an entirely auditory interface. A visual display may be inaccessible because the user is visually-impaired; it may be that the screen on the device is too small (e.g. mobile phones) or non-existent (e.g. the Apple iPod Shuffle); it could be that the user's eyes are occupied with some other task (e.g. driving). However, the development of mainstream auditory interfaces is still in its infancy and there is much work to be done to bring about truly usable applications. The area I address in this thesis is the ease of use of an auditory application, particularly in respect to presenting the user with high-level information so that they have a sense of context.

Many auditory applications launch the user straight into the detailed information and the onus is on the user to keep track of what is going on in the interface. In a graphical interface, the persistence of the display keeps the user immersed in their task, providing peripheral information such as options available, navigation, etc., all of which centres the user and enables them to move efficiently through their task. Audio, on the other hand, is transient, lending itself to providing single pieces of information at a time. This can sometimes result in users swimming in a great deal of information and having a hard time making sense of it all. To address this issue, I looked to how we handle such things in graphical interfaces: when there is a lot of information, we use abstractions or visualisations to organise it. Shneiderman phrases this in terms of *overviews* in his Information Seeking Mantra (Shneiderman, 1996): *overview first, navigate, filter and details-on-demand*.

Using the concept of overviews as a point of departure, I ask the question: what is an overview?

Little work has been done in defining what one is and how it should be used, let alone how one can be applied to auditory interfaces. This is where I focus my research. My first major contribution relates to the nature of overviews and the second relates to their effect on auditory interfaces.

The first contribution is a set of characteristics of overviews in general through an analysis of successful visual overviews (see Chapter 3). Using examples of typical usage, I determined their attributes and characteristics. The overviews studied are varied and incorporate textual and graphical aspects. They were chosen to provide a broad base for the analysis and greater confidence that their similarities describe overviews in an abstract sense. They are: tables of contents, file managers, line graphs, and abstracts. Their differences and commonalities fed into general guidelines as to what makes an overview. Also contributing to the guidelines was knowledge imparted by other researchers into overviews. While they do not focus on the nature of overviews, they do provide arguments for the use of overviews. I reviewed several strands of research, visual and auditory alike, and then re-assessed the information after my survey of common visual overviews. The result was a set of characteristics that overviews should have and they are as follows:

- showing scope,
- quickly guiding,
- showing context,
- describing the structure,
- encouraging exploration and
- providing a historical snapshot.

These guidelines allowed me to approach an evaluation of how auditory overviews work in practice – the second major contribution – with a clearer understanding of the role they play. While most auditory overview research evaluates overviews on their own, my particular focus was on how the inclusion of an overview affects interaction with an auditory interface. The rationale for this decision was motivated by the fact that overviews are not intended to be used in a vacuum: they should be paired with the detailed information they describe. Using a transportation network map application, I examined how usage of an auditory application changes when paired with an overview. Here, the aim of the overview is the display of stations and lines and their interconnections in order to facilitate route-finding. This interface is described in Chapter 4. Participants used the application with and without overviews and solved a series of complex route-finding tasks. I assessed the usage of the application through task completion

metrics, solution correctness and navigation patterns. This experiment showed that, with an overview, participants used significantly different interaction patterns in the auditory application. The implication of the results is that auditory interfaces could be streamlined through usage of an overview.

The organisation of this thesis is as follows: Chapter 2 contains pertinent background research. Chapter 3 describes my survey of overviews; I combine a review of research into overviews with an in-depth analysis of successful visual overviews to develop guidelines about what roles overviews play. Chapter 4 describes the auditory overview developed for this thesis and Chapter 5, its related user study; I analysed navigation patterns and task solving strategies to show how the presence of an overview changes the usage of the application. Chapter 6 contains discussion and concludes the thesis.

Chapter 2

Background

This chapter sets the scene for this thesis and reviews research pertinent to overviews and auditory interfaces. Section 2.1 describes the motivations behind the work and how I came to study this particular aspect of auditory interfaces. Section 2.2 details current research into auditory interfaces and, in particular, auditory overviews.

2.1 Motivations

2.1.1 Why audio?

In the field of Human-Computer Interaction (HCI), researchers and practitioners are constantly looking to improve the user experience. This can entail better support for tasks users wish to complete and making computers easier, more enjoyable and more effective to use. Over the decades, computing paradigms have evolved from paper printouts of results to graphical interfaces and direct manipulation (Eberts, 1994). Likewise, computers have evolved from taking up entire rooms to desktop and mobile computing in the form of laptops, mobile phones and embedded devices. As computer usage expands from the office environment into the wild, screens have become smaller to make them more portable, but it also limits their use. This section describes some of the advantages of using audio as a primary output modality.

Restricted screen space

As a result of diminished screen space, some HCI practitioners and researchers are turning to alternate output modalities. A clear example of this is the iPod Shuffle mp3 player[1] where the screen has been eliminated entirely, enabling the device to be very small and portable. The chosen output modality in this case is audio: it uses VoiceOver[2] to read out the name of the currently playing content and other device information. The device is made to be listened to, making audio the natural choice for interaction.

Eyes-free tasks

Audio is also useful when supporting tasks where the user's eyes are otherwise occupied. As devices are used more on the go, the eyes are more likely to be engaged in other activities. Good examples of this are GPS navigation devices where audio is used to give detailed directions while the screen mimics the road ahead. In this case, it is not safe for the driver to read large amounts of text on the screen but the instructions need to be detailed. Speech output enables exact information to be spoken to the driver, keeping their eyes on the road.

Visual overload

Another case where audio can be an advantage is when the user is receiving too much information visually. The eyes have too much to process and offloading some streams of information to another modality can distribute the load (Moreno & Mayer, 1999). This is generally seen in specialist applications. One area where this has shown some success is for stock traders (Childs, 2005). Often, they will have several screens displaying real-time information. Additionally, they might be conducting business on the phone. Here, non-speech sounds can be used to present information. Audio, being omni-directional, lends itself well to the work environment in question.

Where audio is uniquely suited

In other cases, audio is used where it is uniquely suited to the application area. Audio presents patterns well and human perception of frequency allow us to identify changes in the sound stream (see Bregman, 1994). One case where audio is appropriate is the sonification of seismic data (e.g. Dombois, 2001; Joseph & Lodha, 2002) where tremors map naturally to frequency. This is due to similarities in the haptic and auditory modalities which make them work well together (e.g. Brewster, 2005; Yu & Brewster, 2002). Another case where audio is well-suited to the task is the sonification of Electroencephalogram

---

[1] http://www.apple.com/ipodshuffle/
[2] http://support.apple.com/kb/HT4322

(EEG) data. Hermann *et al.* (2006), who have done much work in the area of bio-medical sonification, sonified EEG data using vocal sounds. Being vocal sounds, they are easily reproduced and this allowed doctors to mimic the sound of an epileptic fit in order to discuss the event with each other.

Accessibility

Another area where audio is heavily used is accessibility for Visually-Impaired (VI) users. With either limited or no access to the graphical display, audio can present the information displayed instead. The two primary solutions for visually-impaired users are braille and auditory output. Audio has the advantage that no specialist equipment is necessary as most computers come with the ability to produce sound, and headphones or speakers are inexpensive and readily available. However, specialist software, called a *screen reader*, is necessary. Screen readers are accessibility software that use text-to-speech, and increasingly non-speech sounds, to display what is on the screen.

2.1.2 Audio for accessibility

The above section describes some of the reasons why audio is attractive. There are myriad foci for research, each with its own challenges. For example, restricted screen space lends itself to alternative display methods and innovative display techniques. However, accessibility involves many of the factors that affect the others. Providing auditory interfaces for VI users, which includes blind users and partially-sighted users who can use large print, has the advantage that the target user group is well-defined. Another appeal is that the justification for using audio is acceptable and straight-forward.

Providing solutions for VI users can be easily generalised for the general population while the inverse is not necessarily true. This is a design concept called inclusive or universal design. A good example of the success of inclusive design – where solutions for special populations can benefit the general market – is OXO Good Grips[3] where the inspiration was the difficulty many arthritic people have with standard kitchen implements (Design Council, 2011).

Screen Readers

There are a variety of screen reader applications to support VI users. What follows is a basic description of several of the more popular ones. The primary output modality is speech however more of them are incorporating non-speech sounds to provide additional information as the application reads out information about the current focus of the cursor. These sounds are typically audio alerts.

---

[3]http://www.oxo.com/s-21-good-grips.aspx

Apple[4] provides VoiceOver[5] with several of their products: OSX 10.6 (Snow Leopard); the iPhone 4 and 3GS[6]; the iPad[7] and the iPod Shuffle[8]. In terms of non-speech sound, VoiceOver uses sound effects, or audio alerts, to indicate interaction effects such as being on the last item in a list or blank space. More complex sound is used in progress bars where the user hears a ticking sound, where the frequency of the ticking represents download speeds.

JAWS (Job Access With Speech) for Windows[9] is a screen reader that works on Microsoft Windows XP and beyond. It includes the facility to add audio alerts to indicate changes in text such as font and formatting. The most recent version provides the Text Analyzer[10] which indicates errors that are not obvious while speaking the text using non-speech sound. These are errors such as extra white space and capitalisation problems.

NVDA (Non-Visual Desktop Access)[11] is a free screen reader for Windows. As an open source project, the project allows for trying out new features to improve accessibility[12]. However, the innovations are currently poorly documented. In terms of non-speech sound, NVDA uses rising tones to indicate progress both in progress bars and other situations, such as loading a page.

Alternatives to auditory display

The main accessibility alternative to audio output for computer usage for VI users is a braille display. In terms of usability, braille displays provide some advantages, such as combining haptic feedback with the braille displayed text (e.g. Luk *et al.* (2006)), and some disadvantages, such as moving the user's hands away from the keyboard or other input device. However, the two main drawbacks, detailed below, are (1) the cost and bulk of the equipment and (2) limited braille literacy.

Equipment for braille   In a thorough survey of assistive technology and their related costs including braille keyboards and displays, Pal *et al.* (2011) state a braille display ranges in cost from $2,000 to $15,000. A typical portable one, such as the Freedom Scientific PAC Mate[13], has 20 to 40 cells – a cell is a braille character – and is about the same size as a QWERTY keyboard. This is in addition to the

---

[4]http://www.apple.com/
[5]http://www.apple.com/accessibility/voiceover/
[6]http://www.apple.com/accessibility/iphone/vision.html
[7]http://support.apple.com/kb/HT4064
[8]http://support.apple.com/kb/HT4046
[9]http://www.freedomscientific.com/products/fs/jaws-product-page.asp
[10]http://www.freedomscientific.com/downloads/jaws/JAWS-whats-new.asp
[11]http://www.nvda-project.org/
[12]http://www.nvda-project.org/wiki/About
[13]http://www.freedomscientific.com/products/fs/pacmate-braille-product-page.asp

standard QWERTY keyboard needed for input. The cost of the display makes it difficult to afford while its bulk makes it more cumbersome to use.

Braille literacy   Current statistics show that less than 10% of VI school age children in America are being taught braille (National Federation of the Blind Jernigan Institute, 2009). There are multiple reasons for this, ranging from lack of resources to changing technologies. However, the impact is that fewer people are proficient in braille, making it a less universal method of understanding text. Another issue is reduced tactile acuity with ageing making it harder for older VI users to continue to use braille. A third issue is the increase of diabetes related blindness (MMWR, 1996) which can be accompanied by reduced tactile acuity (Nakada & Lee Dellon, 1989). With less sensitive two-point discrimination, limited levels of proficiency with braille can be achieved.

Implications for accessibility   In terms of how screen readers and braille displays differ in presentation of computer information, a user's choice to use one over another or in combination is likely to be a matter of personal preference. However, in terms of this thesis, the focus is on audio for several reasons: (1) audio is the cheaper solution, (2) it requires less specialist knowledge than learning or being proficient in braille and (3) it has broader applications than for the VI community.

Improvements needed

As described above, there are certain advantages to choosing audio for providing accessibility to VI users. However, the technology of screen readers has room for improvement relating to usability. Much of the work on assessing usability for users of screen readers focuses on web usability. This has become increasingly important as more computing and distribution of information shifts to the Internet. One of the major complaints regarding web site accessibility is that just because the web site is accessible does not mean that it is usable. One particularly damning study was conducted by Petrie *et al.* (2004) where VI users failed at 47% of their tasks when using a screen reader to access a web page; 45% of the problems encountered were unrelated to web accessibility guidelines and checkpoints[14].

Here are some typical problems that VI users encounter with screen readers:

Filling in forms: Forms on the web are often poorly coded resulting in missed information, or poorly structured forcing the user to switch between reading and editing modes to properly fill them in. (Murphy, 2007; Stockman & Metatla, 2008; Theofanos & Redish, 2003)

---

[14]http://www.w3.org/TR/WCAG10/

What is present: There is often a lack of clarity as to what information is on a page. The primary strategy for finding information on a page is to search for text. This can be hit and miss: if the search term is wrong or the search does not loop, the information might not be found. (Murphy, 2007; Theofanos & Redish, 2003)

Scanning for information: Much like sighted users, VI users will scan pages for information. They listen to the first portion of text and if it sounds unlikely to have what they are looking for, they skip the remaining content. This strategy can inadvertently skip over the object of their search. (Theofanos & Redish, 2003)

Text-to-speech: Comprehension of a page can be compromised by the screen reader mispronouncing words, e.g. "homepage" (hommapage) or "skip to content" (skip to con**tent**). (Theofanos & Redish, 2003)

Familiarity: To effectively use a screen reader, users must be proficient in its features. This can add another level of complexity for the user. (Theofanos & Redish, 2003)

Screen readers not only have problems exposing web pages to VI users: many aspects of standard Graphical User Interface (GUI) applications prove difficult to translate to speech. Navigation, user memory load and improper display of data are all typical problems. Here are a few reported in the literature:

Lengthy navigation: Users will often need to listen to 15-75 items or more in a single navigation sequence. Deep and sparse menu and folder structures are time-consuming and difficult to navigate. (Barnicle, 2000; Hochheiser & Lazar, 2010)

Lack of context: Items, such as menu items, controls or buttons, might be labelled but their meaning only comes from knowing what else is around it. This results in a lot of guesswork on the part of the user. (Barnicle, 2000)

Focus problems: Many errors arise from confusion over what the screen reader is focused on. For example, changing a font in a word processor might fail silently if the focus has come off the text the user wants to manipulate. (Barnicle, 2000)

Memory load: As audio is transient, users can easily forget a piece of information they just looked up

in order to complete a task. To recover, they can either go back and get the information again or note it in a third application. (Parente & Bishop, 2009)

Spreadsheets: The versatility of spreadsheets cause numerous problems. The location of the data is not always evident, charts are not exposed well and the numerical data are presented without anything to help make sense of it. (Doush, 2009; Stockman, 2005)

Learning curve: There is rarely guidance specifically tailored to screen reader usage for applications, rendering the help files difficult to understand and put in practice. VI usage of applications is largely ignored by the developers. More often, it is the screen reader documentation itself that provides some specific guidance on how to use popular applications and it is less extensive than the help files that are written by the developers for sighted users. This makes the learning curve steep and discourages users from straying from what they already know. (Earl, 1999)

Collaboration: Another case that proves difficult is collaborating with sighted users. The users do not always have access to the same actions or the same layout. The order of presentation can be different and features such as spatial layout, which sighted users frequently use as points of reference, are not exposed. This makes it difficult for each to guide the other to the correct place to perform some action. (Stockman & Metatla, 2008)

2.1.3   The case for overviews

Two themes that emerge from the above usability issues are the lack of clarity about what is in the interface and what the context is for individual items being spoken. I posit that overviews can be a solution to these issues. This is backed up by findings from various researchers in the area of auditory interfaces. Murphy (2007) developed a web browser for VI users that attempts to convey spatial information via auditory cues. User feedback specifically requested overviews to aid with gaining familiarity with new pages. In Parente & Bishop (2009), users stated that summaries were important for task efficiency.

In his seminal paper, Shneiderman (1996) advocates the use of overviews as the first step in any info seeking task: *Overview first, zoom and filter, then details-on-demand*. He argues that the overview allows users to conceptualise what is in the data set. He also states that query searches are not sufficient for understanding a collection. Below are multiple examples of applications of this concept in graphical interfaces.

North *et al.* (1996) present an image collection of the human body which allowed quick scanning of all the materials available. Users would mouse over an image of a human body and thumbnails of images available for each body part would show. Feedback reported that it was easier to determine what was in the collection and easier to find what one was looking for. Plaisant *et al.* (1996, 1998a) describe the interface and evaluation of LifeLines, a system including an overview for displaying medical or other personal history, where the overview facilitated identification of interrelationships between distinct records. Hornbæk & Frøkjær (2001) present an overview+detail view of documents where the overview is a thumbnail of the entire document, taking advantage of the size of the headers to describe its geography. The results of their experiment, which compared overview+detail to linear and fish-eye views, indicated that recall and understanding of the documents was better with overview+detail. Users stated a preference for the overview+detail because of easier navigation and the ability to see the document structure. In task efficiency measures, the overview+detail view was slowest but on par with a linear view. I discuss the validity of thumbnails as overviews in Section 3.1.3. In Hornbæk & Frøkjær (2001), thumbnails worked well as the document being described used very obvious features that were still visible at a smaller scale – large headers were still legible. Hornbæk *et al.* (2002) compared usage of zoomable maps with and without an overview. The overview in this case was a thumbnail marked up with the current detail view. Their results show, much like mine in Chapter 4, that overviews did not provide an advantage in terms of task correctness and task completion times were often slower. However, feedback indicated that users preferred the interface with the overview as it was less disorienting. Kustanowitz & Shneiderman (2005) describes a more sophisticated use of thumbnails in an overview of photo collections: a main photo is surrounded by an overview of categorised photos relating to the main one. The evaluation was limited but the feedback was positive and the interface encouraged participants to explore the library in a way that an uncategorised one does not.

The above research reveals several rationales for using and incorporating overviews. Shneiderman (1996) talks about overviews allowing users to conceive of a data set. This is very vague and the studies above provide more specifics. One motif that is touched on often in the literature is that of knowing what is there (Hornbæk & Frøkjær, 2001; Kustanowitz & Shneiderman, 2005; North *et al.*, 1996). The presence of an overview shows the scope of the collection in a way that search queries do not. This directly addresses some of the difficulties VI users have with spreadsheets and web pages I discussed earlier. In order to effectively search for some information, the user needs an idea of what is present.

Another motif is that of orientation and navigation (Hornbæk & Frøkjær, 2001; Hornbæk *et al.*, 2002; North *et al.*, 1996). Most of this evidence comes not from analysing the usage of the interface but from the user feedback. Regardless of whether users performed the tasks more quickly or effectively, they found the presence of the overview reassuring and they felt less disoriented. From their point of view, it was easier to find what they were looking for. Kwasnik (1992), who analysed various types of browsing and navigation, states that browsing in context reduced ambiguity. This is the same concept. This facet of overviews addresses the problems that VI users have with lack of context, scanning for information and lengthy navigation. It may even address focus problems if the overview can more clearly identify where the user is in the interface.

A third motif is discovery, exploration and interrelationships (Kustanowitz & Shneiderman, 2005; Plaisant *et al.*, 1996, 1998a). The overviews made it easier to browse and explore. With LifeLines (Plaisant *et al.*, 1996, 1998a), the overview made it more apparent how various distinct events or records created a pattern. The individual records could then be investigated to get the specifics of the story the pattern tells. This is a similar concept as discussed in Keim (1997) where he advocates that one use of overviews is exploratory: to look for trends and patterns in order to form hypotheses. Similarly, relationships between photos in Kustanowitz & Shneiderman (2005) prompted connections to be explored. This motif of exploration and interrelationships is a more vague one and does not directly address the usability problems that VI users experienced with screen readers. Rather, it pulls the interaction away from purely functional (looking for something) and more towards the experience ("ooh, look at that"). Frustration with the experience is present in all of the screen reader usability problems. The effective support of exploration has the potential to mitigate the impact of the problems I detailed in the Section 2.1.2 [p. 19].

There have been attempts by screen readers to include more summary information. However, it falls quite short of the desired effect. For example, JAWS includes an "overview" of web pages that is triggered by a keyboard command. This feature lists the number of links, headers, forms and frames. I discuss this form of overview later in Section 6.1.1, however suffice to say that in terms of giving the user an idea of the content, this listing of elements provides little more than an indication of how busy the page is. This listing of elements also does not provide any spatial layout – which would aid in collaboration with sighted users (e.g. is there a column layout? Where is the main navigation?) – or any indication of dynamic elements - which remain a source of confusion and are mainly hidden from VI

users until they are activated.

As stated above, VI users have requested overviews or summaries in their auditory interfaces (which go beyond what JAWS has provided as a web page "overview"). Based on the research reviewed here, overviews appear to be a viable solution for providing a more useful experience to VI users. However, what is lacking from the overview research is abstract information about overviews. Shneiderman (1996) speaks generally about overviews but is so general that he recursively qualifies overviews using the term overviews ("Overview: Gain an overview of the entire collection" [p. 2]). Each of the evaluations above bring out certain qualities of overviews but there is not a sense of commonality. The evaluations are in essence not of overviews but of instances of overviews. This conclusion was backed up by my study of overviews for computer desktops (Nickerson & Stockman, 2005). While the focus of the study was the ability of the overviews to convey information, the main finding (unpublished) was that a better understanding of overviews is required before implementing solutions. In order to resolve this issue, Chapter 3 of this thesis focuses on determining more specifically what an overview is (and is not) and what role it plays.

2.2 Auditory interfaces and overviews

Interest in auditory interfaces stemmed from advances in GUIs making computers less accessible to visually impaired users. Interfaces moved from column and row based interfaces to direct manipulation interfaces with graphical widgets. While the older interfaces were relatively easy to make accessible to the visually impaired, the abstract nature of the desktop metaphor posed a challenge.

This interest in accessibility led to a number of research efforts. Starting in the late 1980's, there was research into auditory representation of graphical icons (Blattner *et al.*, 1989). In 1992, interest in auditory display was sufficiently high for the establishment of the International Community for Auditory Display (ICAD)[15] and the first conferences led to the publication of Kramer's seminal tome on auditory display (1994). This was followed by attempts to implement full auditory interfaces for desktop computers. The Mercator Project (Mynatt, 1995) at the Georgia Institute of Technology was an auditory design for X Windows. This interface primarily mapped screen objects into audio using representations for windows, menus and other interface objects. There were also efforts to provide mouse navigation for these interfaces (Edwards, 1988; Savidis *et al.*, 1996). It quickly became evident that there was a need

---

[15]http://www.icad.org

to better understand the auditory modality and sound design. This led to research in non-speech sound (e.g. Brewster, 1998; McGookin & Brewster, 2004a; Walker & Lindsay, 2003).

This section outlines various research that builds towards auditory overviews: auditory perception, sonification, auditory interfaces and finally auditory overviews.

2.2.1 Making non-speech sounds

Auditory perception

For the most part, auditory overview makes use of non-speech sound. Auditory perception describes how we hear and interpret these sounds. The basics of sound are described by physics: audio is a pressure wave that can be described in terms of frequency, phase and amplitude. These describe the physical attributes of sound. The perceptual attributes of sound are more generally described using pitch, timbre, duration, loudness and direction. This section provides a brief set of examples of psychoacoustical research that underpins further auditory display work.

There are many texts that describe the intricacies of auditory perception. One such text is Cook (1999) which goes into detail about the basic anatomy of the ear and how we interpret sound. The field of auditory display builds on such knowledge. Auditory perception, or psychoacoustics, provides a framework of how the pressure waves perceived by our brains turn into information.

Auditory Scene Analysis (ASA) (Bregman, 1994) is a subset of psychoacoustics that looks to understand how we interpret not only certain sound combinations but also our entire sound environment. It is based on the theory that we separate our environment into streams using various aspects of change and difference in pitch, timbre, direction, etc. An example of this is the way time can separate sounds or meld them into a stream: a large silence makes events segregate while sounds closer together in time are associated more closely. This effect is also seen in differences in pitch. A series of alternating high/low tones with similar frequencies form a sound stream while the further apart the frequencies, the more the sounds are perceived as disparate. In other words, similarity and proximity of sounds in time and frequency lead to them being associated together.

Another example of psychoacoustic work is Kubovy & van Valkenburg (2001) where the authors attempt to glean what constitutes a "auditory object" and what its primary and secondary characteristics are. Their findings propose that while in visual objecthood, we are most interested in where the object is; when we listen, we are mostly interested in the source itself. In other words, visually, we focus on 'where' while auditorily we focus on the 'what.'

Auditory interface artefacts

The above examples of research are indicative of the type of knowledge that props up work in auditory interfaces. The more we understand how human auditory perception works, the better we can display information using non-speech sound. However, knowing how we interpret sounds does not guarantee that directly applying the principles of psychoacoustics results in successful interfaces. What follows is a review of some research where the application of psychoacoustics is tested in practice to create building blocks for auditory interfaces.

There have been several attempts to re-create graphical interfaces in audio. One of the key components that received much attention is icons. Icons essentially provide a branding for applications and actions; the idea is that the visual pattern is immediately recognised and associated with the action or software. The two major approaches to creating an auditory equivalent are *auditory icons* and *earcons*. The following sections describe these two approaches and their basis in psychoacoustics. I also include a section on spearcons, a more recent approach.

Auditory icons   Auditory icons come from the tradition of ecological sounds and everyday listening. Gaver (1988) introduced the concept in a paper that formed part of his PhD thesis. It was his opinion that using the attributes of sound (pitch, loudness, etc) was not as powerful as using sounds that listeners are familiar with. Users already know how to interpret these sounds and Gaver posits that various attributes of the source of the sound can convey additional information. For example, a bouncing ball could indicate the arrival of some object and a heavier sounding ball could indicate a larger object. He is, however, quick to point out that this does not necessarily mean that users will grasp the intended message (Gaver, 1988, p. 88).

Some of these concepts were tested in a cola factory simulation (ARKola) (Gaver, 1991; Gaver *et al.*, 1991) however the work never reached the point of altering sound source attributes to convey additional information. Nevertheless, Gaver showed that applying real world sounds with appropriate mappings was effective and had a small learning curve. Participants in his study, which was based on collaborative management of the cola factory, used both individual sounds to diagnose problems and all the sounds together as reassurance that the factory was running well.

The disadvantage of using everyday sounds is that mappings from events to sounds must be tight. In some cases, there may not be an appropriate metaphor. Also, certain sound sources command more attention than others, creating problems of mis-matched urgency. A third caveat is that auditory icons

are best suited to alerts. Alerts are key to providing real-time feedback about events however this is only a small subset of the types of information one would want to expose in a purely audio interface.

Work related to auditory icons has mostly moved into that of ecological soundscapes. There is little that specifically looks at how auditory icons can be applied more effectively. Gaver himself continued his work by looking at algorithms to more effectively synthesise real-world sounds (Beaudouin-Lafon & Gaver, 1994; Gaver, 1993a,b). One project that aimed to take the use of auditory icons and real-world sounds further was *Nomadic Radio*: Sawhney & Schmandt (1999, 2000) developed a purely audio wearable computing environment and made use of a flowing water metaphor to indicate status and to notify of incoming messages and other content. The sound was modulated depending on the urgency and length of the messages. The majority of feedback was provided in speech, however the auditory cues gave users sufficient warning to switch their attention or to stop a message from being played. Another project of interest is Gutwin *et al.* (2011) who evaluated varying the parameters of the sound of chalk to indicate realistic usage in a collaborative workspace. The sounds accurately indicated speed of drawing, pressure and shape in the shared environment.

Earcons   An alternative to auditory icons is earcons (Blattner *et al.*, 1989). Where auditory icons are like a picture, earcons are like an abstract pattern. One can think of earcons as organised sounds that are carefully constructed to create quick and informative messages for use in an auditory interface. Blattner *et al.* lay out a set of guidelines for the development of earcons. The basis of an earcon is a *motive* which has a fixed pitch and rhythm. The motive can then be extended to create hierarchies and families of earcons by making changes in timbre, register and dynamics. This allows compound earcons to be created by combining, for example, an earcon that describes an object and an earcon that describes an action. The result is an earcon that describes that action performed on that object.

While Blattner *et al.* (1989)'s work was theoretical, Brewster *et al.* (1993) and Brewster (1994) addressed their suitability for auditory interfaces and produced guidelines for their creation in practice. Brewster (1994); Brewster & Edwards (1995) explored the possibility of parallel earcons rather than serial ones as proposed by Blattner *et al.*. Instead of lengthy compound earcons, the earcons were played at the same time, reducing the amount of time it takes to present an earcon that, for example, delete a file (*delete* earcon played with *file* earcon). One of the major outcomes of this work was showing that earcons could be effectively used to convey information in an interface. Another was a revision of how to create earcons: while Blattner *et al.* (1989) recommended pitch as a primary component of earcons,

Brewster recommends that timbre play a larger role. Additionally, major differences were more effective than subtle ones, e.g. changes in register should be of several octaves. Rhythm was still an important factor and the recommendation is that varying the number of notes for each earcon makes them more recognisable. These concepts were further tested in the context of menu hierarchies and telephone-based systems (Brewster, 1997; Brewster *et al.*, 1996). The experiments showed that recall rate in context and recognition of new earcons were good. Brewster (1998) went on to show that using hierarchical earcons combined with compound earcons increased their effectiveness. The next question about earcons was how to put them to use when the interface becomes more complex. With the view to impart more information, McGookin (2004) explored the possibility of concurrently presenting unrelated earcons. The research indicates that recognition rates are lower, however they are improved by spatialisation (see also McGookin & Brewster, 2004a), more disparate use of timbres and staggered onsets. These changes were directly inspired by ASA (McGookin & Brewster, 2004b). Examples of other studies that showed benefits of earcons in practice include Spiliotopoulos *et al.* (2009); Yu & Liu (2010b) where earcons were applied to a sonically enhanced menu navigation system in a mobile context and to a Text-to-Speech (TTS) data table system.

Earcons have benefited from more extensive study than auditory icons. There are several potential reasons for this: they are easier to synthesise; there is a greater range of possible sounds; and there is no need to create logical mappings between real-world sounds and their meaning in the interface. This makes earcons more versatile and attractive to developers. The research shows that they are effective and the subsequent design knowledge is extensive. However, where auditory icons have a clear advantage is in learnability (Garzonis *et al.*, 2009). It takes proper training to learn the significance of the abstract sounds of earcons (Brewster, 1997). An open question is how the learnability and recall is affected by long-term usage – most experiments involved limited and short-term use. With prolonged use, the difficulty of recall could be mitigated by familiarity, making the versatility of earcons all the more attractive.

Spearcons Spearcons (Walker *et al.*, 2006) are a relatively recent innovation. They were developed to quickly display large lists of text (such as a list of contacts or a menu) and address some of the learnability issues that affect earcons. The basic premise is that, when browsing long lists, the speech is sped up to the point of not being recognisable as speech however the sound can be interpreted based on the sound and shape of the spearcon. They have been shown to be faster than earcons or auditory

icons (Walker *et al.*, 2006), to have higher levels of accuracy (Walker *et al.*, 2006), to be easier to learn (Palladino & Walker, 2007), to be easier to create (Palladino & Walker, 2007) and to improve efficiency of TTS-only systems (Palladino & Walker, 2008; Spiliotopoulos *et al.*, 2009).

Conversely, Langlois *et al.* (2010) showed that spearcons were not appreciated compared to auditory icons and earcons. They did not provide the richer hierarchical information that the other two could. This may be because of lack of familiarity or that spearcons seem to work better for lists rather than hierarchies. The latter is supported by Yu & Liu (2010a) where they augmented a menu system with spearcons for menu items and auditory icons and earcons for the hierarchy. This hybrid approach showed that sonically-augmented menus on mobile devices were easier and faster to use. This was confirmed in another experiment where mobile devices were used while walking (Yu & Liu, 2010b).

Spearcons, compared to the other two, show more clearly their strengths and weaknesses. They are shorter and relate directly to the word they represent. This makes them more efficient – less listening on the part of the user – and easier to associate correctly. The fact that they relate directly to a word does pose some problems: localisation of an application would require a whole different set of spearcons and similar sounding words reduce their power. Murphy *et al.* (2010) provide additional guidelines to address the latter problem, advocating varying length words that also contain different letters. As the research in the previous paragraph shows (Langlois *et al.*, 2010; Yu & Liu, 2010a,b), spearcons are not always suitable. It is important to note that they were intended for lists of words (Palladino & Walker, 2007; Walker *et al.*, 2006) so this is not surprising. They can however provide benefits when used in conjunction with other approaches (Yu & Liu, 2010a,b).

Sonification

The previous section discusses auditory interface artefacts that can be used to convey information similar to an icon in a GUI. Each has its strengths, weaknesses and appropriate uses. This section looks at a different aspect of auditory interfaces: the representation of larger sets of information. Sonification is a blanket term for displaying information in sound and it can be seen as the auditory equivalent of visualisation (Hermann, 2008). Sonification is a complete representation but it is not quite an auditory interface. Some examples of sonification are auditory graphs (see the 1st International Symposium on Auditory Graphs[16] in 2005) and the display of scientific data. This section covers some of the techniques used in sonification.

---

[16] http://sonify.psych.gatech.edu/ags2005/

Audification   Audification takes data, usually a time-series, and scales the values to the audible frequency range. It is a simple approach so there is little work on the technique itself but more in its usefulness. One application area is seismology where there is a similarity between the waves created by tremors and sound waves. Dombois (2001) showed that audification can expose the unique sound of a region and the type of plate movement; he also found it was easier to separate the signal from the noise. Another area where audification has been tested is helicopter flight data where engineers attempt to troubleshoot prototypes. Pauletto & Hunt (2005) compared the analysis of visual spectrograms to that of audification. They showed that audification could identify problems to the same extent as the spectrograms, especially when identifying noise and power signal levels. Audification has also been studied in the domain of statistics. Frauenberger *et al.* (2007) investigated audification's ability to expose kurtosis (sharply peaked) and skew (asymmetry) in data distributions. While kurtosis could be identified, skewness was not.

While Pauletto & Hunt (2005) and Frauenberger *et al.* (2007) do not discuss guidelines on how to audify time-series, they do both mention difficulties in ensuring that the data is scaled appropriately to be audible. Another common theme is that audification is better at representing data that has similar properties to sound (Dombois, 2001; Frauenberger *et al.*, 2007; Pauletto & Hunt, 2005). For example, Worrall (2009) used audification to present stock market data but ultimately found that pre-processing the data to represent deviation in the data from a central frequency was an effective sonification. In essence, audification can be quite simple to implement when the raw data is a time-series and there is only one dimension to analyse. Its suitability depends on how similar the raw data is to sound waves – in those cases, its performance is quite similar to a spectrogram.

Parameter-mapping   Parameter-mapping is a sonification technique where multi-dimensional data is presented in audio by mapping dimensions to sound attributes such as pitch, timbre, loudness, etc. By choosing correct mappings, characteristics and patterns in the audio expose areas of interest in the data. It is currently one of the most wide-spread approaches due to its ability to handle multi-variate data and is used most often in the sonification of scientific data.

Parameter-mapping sonification has wide-spread uses as the technique is generic. Some examples of application areas are the sonification of seismic data, gestures/movements, EEG data and images. Joseph & Lodha (2002) examined the sonification of seismic data using parameter-mapping. They use various musical qualities to identify areas of interest such as fault exploration (breaks and discontinuities in rock

layers) and layer consistency (stratification in the rock layers). Turning to gesture, Fox *et al.* (2005) developed *SoniMime*, a system for sonifying hand movements. It uses the audio output, human vowel-like sounds, to help users reproduce a gesture. This is similar to Hermann *et al.*'s work on EEG signals, where they map seizure data linearly to vocal sounds. The result is patterns that can be reproduced by humans such as 'oiii-oiii-oiii-oiii' or 'eyaaa eyar eyar eyar.' In Hermann *et al.* (2008), they refine the technique using an algorithmic mapping, kernel regression, which smoothes the interpolation between the various dimensions being represented. In yet another application area, Yeo & Berger (2005) present a framework for sonifying images. The sonification is primarily based on scanning and maps various pixel values to sound.

The difficulty of parameter-mapping is getting the mappings right as well as mitigating problems such as masking and frequency/loudness effects when multiple parameters interact. Walker & Lane (2001) discuss the need to understand mappings properly - what to map to what, the polarity of the mapping and scaling of the mapping. This makes the technique difficult to use initially and Walker & Lane presented an online database of mappings to aid designers in choosing how to sonify data based on its type and the purpose of analysis. The issue is also addressed by Hermann *et al.* (2007) who present a system that self-optimises the ranges of the mappings to better identify areas of interest in the data. The user then can evaluate various iterations of the optimised sonifications to fine-tune them rather than being distracted by having to implement the optimisation her/himself.

Model-based sonification   Model-based sonification (MBS) is an interactive sonification similar to parameter-mapping. It is primarily for presenting large data sets and is based on representing data using a physical model which is then sonified. Unlike the previous two techniques, the sonification is silent until the model is excited. This allows users to interact with the data from various angles and explore the effect of their interaction to learn about the data set.

MBS is the product of Hermann's PhD thesis (2002) and was first introduced in Hermann & Ritter (1999). The motivation was to take advantage of the fact that the human auditory system is sensitive to change and dynamics encoded into a sonification model could allow for greater interpretation of the contents of a data set. Hermann (2002) presents several different sonification models using the dynamics from various advanced physical systems, such as spatial scanning of data through echolocation (data sonograms). The proposed application areas are varied and range from analysis of annotated transcripts of therapeutic sessions – allowing users to follow the progress of a session – to multi-channel image

data (stacked 2d images) – allowing users to identify frequent or marker patterns.

Hermann & Ritter (2004) introduced the concept of a tangible interface to a data set. The model in question is a particle trajectory model and the interactive device is a ball which users can squeeze and shake. This type of excitation of the system feels natural and the effect is to move the data around to listen to how the excitation motion exposes the state of the data. In essence, the data becomes an instrument and the different sounds teach the user about how the data is interconnected (Hermann & Ritter, 2005). This same concept is applied in Milczynski *et al.* (2006) where they use a malleable user interface, called *Elastable*, to explore the properties of iris flowers and the chemical properties of glass [17]. These same data sets were used in Bovermann *et al.* (2006) to explore the usage of tangible interface objects to explore a virtual spatial representation of the data sets. In more novel uses of MBS, Eslambolchilar & Murray-Smith (2008) applied it to navigation of text documents and in Nickerson & Hermann (2008), I applied it to grid-based games. Eslambolchilar & Murray-Smith (2008) aimed to use accelerometer data to sonify interaction with a text document to facilitate its exploration. In Nickerson & Hermann (2008), a sonification model provides secondary auditory information to auditory grid-based games (see Appendix F). The user study showed that in the case of auditory Connect Four, players were able to successfully use the model to strategise.

The above cited research shows that MBS is a versatile technique for sonifying complex multidimensional data sets. However, while many application areas have been explored, there are few user studies documenting its usability. Hermann & Ritter (2004) claims that the models take advantage of everyday listening skills and thus provide sonifications that are already familiar to listeners from interaction with real-world objects. One of the reasons for the lack of user studies is the highly specialist domains to which MBS has been applied. Many involve scientific data where MBS is attempting to assist in data exploration not well suited to the visual domain. Thus, there is a lack of extensive users groups with the right domain knowledge. However, Eslambolchilar & Murray-Smith (2008) and my work in Nickerson & Hermann (2008) show that MBS can be used in more general user interfaces.

2.2.2   Auditory interfaces

The previous section reviews some of the techniques and approaches for making effective use of nonspeech sound. The interface artefacts are building blocks for auditory interfaces while sonification shows

---

[17]"UCI repository of machine learning databases,"
ftp://ftp.ics.uci.edu/pub/machine-learning-databases

what we might be able to express in an auditory interface. This section moves on to auditory interfaces themselves. The majority of auditory interface work focuses on a particular application area. I will first address more general solutions then review interfaces for auditory maps/navigation, web browsing and gaming.

General work on auditory interfaces

The Mercator Project was a major attempt to make GUIs more accessible (Edwards & Mynatt, 1994; Edwards *et al.*, 1993; Mynatt, 1995; Mynatt & Edwards, 1992; Mynatt & Weber, 1994). The auditory solution was implemented on the X11 window manager [18] for Unix. The novelty of the project was that they were not trying to make a better screen reader but rather to provide a similar experience as GUIs do to sighted users. It involved using the cursor keys to navigate the hierarchy of interface widgets.

A number of other researchers addressed various components which would have been necessary for such an all-encompassing auditory interface. Edwards (1988); Jagdish *et al.* (2008); Pitt & Edwards (1991); Winberg & Hellstrom (2003) all researched auditory usage of a mouse. The most common approach is spatialisation of the interface using stereo output or Head Related Transfer Functions (HRTFs). Winberg & Hellstrom (2003) tested a system to enable drag and drop using stylus interaction which showed promising results. One important finding was the importance of individualisation not only in the HRTF but also in the sounds used and in support of different tasks. These attempts to provide mouse access to interfaces do not appear to have matured to the point of being integrated into an application to test their usability in practice. One reason for this might be that ultimately, in order to represent an interface to VI users, it is not important *where* an object is but rather that the structure of the interface be accessible and that the functionality be exposed (Murphy, 2007).

Another approach contemporary to the Mercator Project was *Emacspeak*[19] (Raman, 1996a,b). It is a speech-based interface that worked only in the context of Emacs[20], a text editor for Unix that also enables usage of applications such as a calendar, an email reader, a file manager, etc. Its novelty is that it does not *read* the screen but rather makes the application speak. Emacspeak is still in development today[21] but its popularity is limited by the fact that Emacs is rather esoteric and not widely used in our world of social networking and business documents. A more recent attempt is Clique (Parente, 2005)

---

[18]http://www.x.org/wiki/
[19]http://emacspeak.sourceforge.net/
[20]http://www.gnu.org/software/emacs/
[21]Latest release is version 34.0, May 2011
http://emacspeak.sourceforge.net/releases/release-34.0.html

which uses a conversation metaphor – using up to four spatialised concurrent voices – and task-based interaction to provide auditory access to GUIs. The main output modality is speech, however non-speech sound provides secondary information about the interface. Parente & Clippingdale (2006) show that certain aspects of their work can improve usability (concurrency and effortless search features). Such improvements can be integrated into current screen readers without major disruption while task-based interaction requires a change in paradigm. Their work is continuing to strengthen the usability of the task-based interaction by studying how a human would "screen read" an interface to a VI user (Parente & Bishop, 2009).

To date, general approaches to auditory interfaces focus on providing detailed interaction. In Mercator, it was telling the user about where s/he was in the hierarchy of interface objects (windows, menus, buttons, etc.); in other work (Edwards, 1988; Jagdish *et al.*, 2008; Pitt & Edwards, 1991), it was telling the user where the mouse was on the screen; in Clique, it was about guiding the user through a task. Knowing where the mouse is and what action can currently be executed only becomes useful when the interface, and therefore the context, is already known. An expert user could take advantage of the information provided, however a novice user, or one interacting with dynamic information, would need to map out the rest of the screen in order to understand what their current options mean. This viewpoint is backed up by findings from Murphy (2007); Parente & Bishop (2009) where users express a desire for overviews and summaries.

Auditory maps and navigation

Auditory mapping and navigation are an application area of interest to both VI and sighted users. For the former, it is about mobility and getting around independently; for the latter, the interest stems from eyes-free interaction. Auditory maps allow users to familiarise themselves with an unknown area while auditory navigation guides the user on the ground.

The main goal of auditory maps is to present what is the immediate area of a particular location. One example of map exploration is Pielot *et al.* (2007) where they used a tangible user interface (movable marker on a table top interface) to allow users to explore a virtual space. Users of the system made heavy use of rotation to determine orientation and showed that a tangible device could improve usability of an auditory map. Heuten *et al.* (2007); Milne *et al.* (2011) explored a similar approach but with pen and graphics tablet systems. Heuten *et al.* (2007) used a torch metaphor whereby users could only hear objects within a particular radius as too many objects proved confusing. Parente & Bishop (2003)

created audio-tactile maps using spatialised auditory icons and consumer-grade force feedback devices such as trackballs and joysticks. All four interfaces allowed users to gain some familiarity with the area. However, the knowledge the map provided was never tested as the experiments were more about proving the concept.

Like maps, auditory navigation displays the immediate area but with the added goal of guiding a user through the space. Davies *et al.* (2006, 2007) proposed a system that would extend the capabilities of a white cane. It was based on a hybrid approach of sonar and video information where earcons would represent distance to obstacles, auditory icons would represent objects and tempo would indicate speed of movement. Walker & Lindsay (2003); Walker *et al.* (2006) explored the concept of using beacons to lead users to landmarks. Their studies, conducted in a virtual space navigated with a joystick, involved specifying the best sound beacons to use (noise or sonar ping) and the best sizing for beacon radii (1.5m). Marston *et al.* (2006) took their guidance system into the wild and had participants navigate both a city and park route using GPS for location and beacons for guidance. The auditory display used only simple beeps to augment a speech-based system that announced distance and angle to the next landmark. On average, the system was able to guide users through the routes with minimal deviation from the optimal routes. Magnusson *et al.* (2010) used a similar approach however the work is still in initial stages. Crease & Lau (2004) tested using sounds to guide participants using a mobile device and it proved as effective a technique as using visual cues. Jones *et al.* (2008) adapted the sound beacon idea by using volume and panning of music to guide participants through routes. The intended users are sighted and they experimented in both a virtual environment and in the wild. For the most part, participants completed their journeys taking 25% more time and covering 25% more ground compared to an expert who already knows the route. Similar approaches are described in Strachan *et al.* (2005).

With auditory navigation, the sonification is directly supporting an active task. It is therefore appropriate that the interface focus on the immediate details and on augmenting what the user gets directly from the environment. On the other hand, auditory maps could benefit from more overview information. The only evidence of overviews was the virtual walk-through provided by Heuten *et al.* (2007). Like with the early attempts to provide access to GUIs, the user is provided with the detail information and it is up to her/him to formulate an overview based on her/his exploration.

Accessible web browsing

Non-visual web page access is dominated by hybrid audio/haptic solutions. One of the first to propose such a solution was Petrucci *et al.* (2000). They used non-speech spatialised sound and a haptic touch screen to provide access to page layout. The system was never tested and it is not clear whether it was fully implemented either. Donker *et al.* (2002) focused on an audio-only approach using a torch metaphor to describe a section of a page at a time. Unlike in auditory maps (e.g. Heuten *et al.*, 2007), the metaphor does not work well. This is mostly because the spatial proximity of items does not necessarily indicate a relationship between them. Their implementation led to slower interaction and less complete discovery of items on the page when compared to screen reader performance. As a result, the authors proposed an approach based on document structure rather than layout. Another finding from Donker *et al.* (2002) was that users were overloaded by the amount of audio information. Yu *et al.* (2005) used a similar proximity approach yet saw no such effect, possibly because the addition of haptics allowed the load to be better distributed. Murphy (2007) built upon this work and investigated the use of semiotics in sound design of auditory interfaces. One of her main findings, as mentioned in Section 2.1.3, was the need for an overview. The original intent was to provide basic layout information. However, it emerged that what was required was a "macro" view that describes the page content prior to interaction. She discovered that users find the hierarchical structure as important as the layout. I address this concept in Chapter 3.

A further approach is that of Sato *et al.* (2011) where they provide hints via whispers. This is similar to Parente (2008) where he employs four voices, spatially located, to indicate different types of information. In Sato *et al.* (2011), the whispers provide contextual information (e.g. "this is a popular link" or "the cost is..."). In user tests, the presence of the whispers did not affect interaction time or task completion however user feedback was higher, indicating that the effect was one of confidence. This was backed up by more direct navigation patterns (less going back and rechecking of information). While the whispers did not constitute an overview per se, they did provide context, allowing the users to gain a more meaningful impression of the web page. I discuss this type of information further in Section 3.1.3.

Audio gaming

Auditory gaming is a special case of interaction where usability and engagement need to be balanced to keep the player interested. Interest in auditory games stems not only from the VI community but also

from mobile games where audio can play a greater role (e.g. Collins *et al.*, 2010; Liljedahl *et al.*, 2007; Papworth, 2010; Paterson *et al.*, 2010). Visual games often include maps and other features to help orient the player. These, or their equivalents, are often lacking in audio games. Instead many choose to focus on the rendering of 3D environments for audio gaming (e.g. Cowan & Kapralos, 2008; Röber & Masuch, 2004; Sonnek & Rice, 2010). For those who do address the games themselves, most focus on *reactionary* games, *exploratory* games and *teaching* games.

Friberg & Gärdenfors (2004) address two types of auditory games: reactionary and exploratory. *Mudsplat* relies on sound localisation skills: players try to counter mud being slung at them. As a game based on reaction, it contains no overview. *Tim's Journey* is an exploratory game where players solve a mystery by exploring an island. One of the key aspects of the game is discovering the environment and thus, the authors argue, it is inappropriate to provide an overview. However, it does provide orientation via a compass and a way to reduce the audio environment to only salient (i.e. sounds that can be interacted with). They acknowledge that their approach results in slow, exploratory game play as the user needs to generate her/his own overview of her/his environment.

This type of slow interaction is typical of exploratory games. *Shades of Doom*[22] is a popular first-person shooter based on *Doom*[23]. An extensive set of keyboard commands permits easy game play and the sounds guide players through the environment. As each object in the game has a unique sound, these help provide reference points. There is also the ability to create Braille-ready maps and to use the "environment analyzer computer" to navigate when players get lost. However, it is easy to get confused and there exists extensive game help generated by players[24]. Similarly, *TextSL* (Oktay & Folmer, 2010) provided screen reader access to *Second Life*[25] and found that VI game play was inherently slower in both interaction and navigation. As part of their implementation, the authors included a summariser to keep players from being overwhelmed by the number of objects in their vicinity. These are steps in the right direction – however they do not provide the ease of use of their visual counterparts.

Trewin *et al.* (2009) provided accessibility options for *PowerUp*[26], an educational game that teaches how to apply maths and science to environmental problems. The adaptation for VI users included speech output for dialogues and extra sound effects that indicated direction of travel (via footsteps)

---

[22]http://www.gmagames.com/sod.html
[23]http://www.idsoftware.com/
[24]http://audiogames.net/page.php?pagefile=cheats
[25]http://www.secondlife.com
[26]http://www.powerupthegame.org/

and object types (via more realistic sounds). They also implemented a set of special commands such as *auto-aim* and *auto-walk* to facilitate game play. Like in *TextSL*, there is a summary feature to list nearby interactive objects. In their evaluation, only one of their two screen reader participants made use of all the accessibility options which may be explained based on computer experience. All the VI participants expressed a preference for speech over non-speech for receiving information, though the authors acknowledged that the non-speech sounds may prove more desired by advanced users. This sort of feedback implies, as I discuss in Chapter 4, that speech is more appropriate for detail while non-speech can provide overviews.

Like Trewin *et al.* (2009), Morelli *et al.* (2010); Sánchez *et al.* (2010a,b) explore the usage of games for educational purposes. In this case, the focus is on improving the skills of VI children to better prepare them for real environments. Sánchez *et al.* (2010a,b) teach orientation and mobility skills. In their game, players learn to navigate a virtual environment and then apply their skills to the real world. Sánchez *et al.* (2010b) use keyboard interaction and participants were allowed exploration time as well as game time. Sánchez *et al.* noted that the strategy game (where the aim is to find jewels, move them outside and avoid monsters) led to more engrossed play and more complete exploration of the environment. In Sánchez *et al.* (2010a), interaction with the virtual environment was through a haptic "carpet" where players stood in the middle and moved their avatar by stepping in the direction they wanted to go. The players were tested on various orientation and mobility indicators pre- and post- play. The major effect of the navigation tasks was increased tempo-spatial orientation. That is, the users had a better sense of their location and orientation in space in terms of what was in front, behind, to the side and how much space lay between them. The sounds and speech in the environment were designed to mimic the real world – in that sense, there was no built in overview. However, the participants were introduced to a tactile model to familiarise them with the environment. This model played two roles: to help players better interpret the audio environment and to provide them with a sense of the layout prior to use. They did not have access to the model during interaction, nor did they have models for all they spaces they explored. This tactile model can be seen as an overview and prior work by the first author showed that physical models aided in the understanding of their associated virtual environments.

Morelli *et al.* (2010) looked at a different aspect of VI adaptation: physical exercise. Many physical activities are difficult for VI people to engage in and the authors sought to address how to help VI children achieve the health benefits of exercise through games. They adapted *Wii Tennis*[27] to replace

---

[27]http://wiisports.nintendo.com/games_section/

the visual cues with auditory ones; vibro-tactile cues were augmented to include the ball bouncing; visual cues such as ball speed and distance were ignored as it proved too distracting for players. They compared audio-only with audio/tactile VI Tennis. Both were equally playable though audio-tactile VI Tennis was more engaging. Both provided the same level of physical exertion though it falls short of what is sufficient for children (it is more suitable for adult activity requirements). As a reaction-based game, there were no overviews included.

Most of the examples of audio gaming above focus on what audio is good at: describing the "what" and identifying its approach. These games tend not to include overview as the challenge of the game is the player's reaction time. This is true of other rhythm-based games such as *Finger Dance* (Miller *et al.*, 2007), *Rock Vibe* (Allman *et al.*, 2009) and *Blind Hero* (Yuan & Folmer, 2008). Where there could be some sort of "look ahead" or other overview (Morelli *et al.*, 2010; Yuan & Folmer, 2008), they are deemed too confusing. In games for orientation and mobility, the goal is to fidelity to the real world environment, so again an overview is not provided. However, Sánchez *et al.* (2010a,b) provide a tactile model that fulfils the same role as an overview prior to interaction. In exploratory games (Friberg & Gärdenfors, 2004; Oktay & Folmer, 2010, and *Shades of Doom*), where an overview could be a real benefit, overviews take the form of summaries that only expose the immediate area. These attempts, however, acknowledge the need for overview-type support for game play. In other words, audio gaming remains focused on the immediate, the detail. This leaves VI players shouldering the burden of creating mental models that sighted players would get through maps or other features.

2.2.3   Auditory overviews

Auditory overviews are not currently a major part of auditory interfaces. The previous sections highlight their absence and where they could benefit. There are, however, a few niche areas where auditory overviews have taken hold and have been explored in terms of their creation and usability. These areas are information-seeking, graphs, tabular data, computer source code and web browsing. The following is a brief summary of research that will be discussed in detail in Section 3.1.1.

Auditory overviews in information-seeking

Information-seeking is rather like data exploration in that vast amounts of data are available and it is difficult to absorb it all. Unlike data exploration, the user is not interested in merely the detail but rather in what it imparts (hence information instead of data). Information-seeking implies a large data set

and Shneiderman (1996) championed the need for overviews to help make sense of it all. In general, visualisations and visual overviews fulfil this need. Details of some of his work – along with that of his colleague Plaisant – are discussed in Section 3.1.1. In terms of auditory overview work, the mantle was taken up by Zhao (2006) who adapted the Information-Seeking Mantra (Shneiderman, 1996) – *overview first, navigate, filter and details-on-demand* – to auditory interfaces. She termed this *Auditory Information-Seeking Actions (AISA)*, using the term *gist* to indicate overviews. This was to avoid visual nomenclature.

After some preliminary work to test the auditory adaptation of the info-seeking mantra to determine successful ways to present *gists* (Zhao *et al.*, 2004), Zhao (2006) settled on a sweep approach. The main focus of her work was geo-referenced data, specifically U.S. census data by state. Various types of information could be presented, such as total population, population by age range, etc. The data was visually displayed on a choropleth map where values were grouped into five ranges. The *gist*, or auditory overview, swept through the states horizontally and a spatialised tone played the value range of each state. This method could be repeated for subsections of the map to create sub-*gists*. Zhao (2006); Zhao *et al.* (2008) found that the auditory overview successfully displayed patterns in the data and encouraged exploration of areas of interest. They propose that this method can be extended to other forms of data, using scatter plots as an example.

Since the *gist* was part of a greater whole (exploration of the data-set), there are few details on the strengths and weakness of the sweep method. In Zhao *et al.* (2004), a data table approach was disregarded as participants preferred the spatial layout that correlated to the configuration of the map and ultimately thought that they would become more proficient at such an approach. This implies that the shape of the overview and its relation to the underlying data is an important aspect of overviews. This theme is mirrored in the creation overviews of node/edge graphs (Brown *et al.*, 2006), which I discuss in the following section.

Auditory graphs

Audio graphs, like their visual counterparts, are overviews of data. They have received a lot of attention due to their prevalent usage: visual graphs are integrated in to spreadsheet software and often appear as graphics in news articles. Accessible versions of audio graphs are for the most part line graphs; I address the numerous guidelines for their generation below and discuss them further in Section 3.1.2. I will also discuss some work relating to edge/node graphs.

Line graphs The guidelines for audio line graphs cover many aspects of their generation: what MIDI note ranges to use (Brown *et al.*, 2003), mapping *y*-values to pitch (Brown *et al.*, 2003; Flowers, 2005), how to use parallel and serial representations for multi-line graphs Brown *et al.* (2003), etc. The general consensus is that continuous presentation (a single varying sound as opposed to discrete sounds) is best for giving an overall impression (Flowers, 2005; Harrar & Stockman, 2007) however discrete presentation of tick marks provide better access to context and exact values (Harrar & Stockman, 2007; Nees & Walker, 2008; Smith & Walker, 2002), Another recommendation is to use musical timbres rather than pure tones (Brown *et al.*, 2003; Childs, 2005). Panning and stereo output can also help in their interpretation (Brown *et al.*, 2003; Ramloll *et al.*, 2000).

While we know much about how to generate auditory line graphs and that they can successfully communicate the gist of the underlying data (Brown *et al.*, 2003; Childs, 2005; Flowers, 2005; Harrar & Stockman, 2007), their interpretation is still a challenge for users. Nees & Walker (2008) compiled qualitative data about strategies that participants use to understand auditory graphs. The descriptions of the strategies show that users think a lot about how to succeed in interpreting them correctly (e.g. "I counted the snare beats on my fingers and remembered the finger I lifted at the correct time (thumb at 8am, index finger at 9, etc.)."). Most people used a combination of strategies and the most common are using context and counting. Using context means comparing notes to each other: is a tone higher or lower than another? Is the tone ascending or descending? Counting refers to counting out time to establish correct separation between values. These findings corroborate what Flowers (2005) advocates and traces back to what the human auditory system is good at computing. This indicates that the challenge with auditory graphs is not in understanding what they convey but rather in familiarity with the task.

Edge/node graphs Another type of graph that is less studied is the edge/node graph. The reason it presents such a challenge is its spatial layout which is difficult to translate to the auditory modality. As discussed in Kubovy & van Valkenburg (2001), we are much more adept at the audio *what* than the *where*. Cohen *et al.* (2006) describe an application for sonifying edge/node graphs where they concentrate on the details level: users navigate to nodes and listen to a sonification of nodes and edges connected to that node. Users were successfully able to interpret the graphs, however the need for overviews is made clear when the authors state it takes approximately three minutes to understand a four-node graph. Even without empirical evidence about sighted performance with edge/node graphs, three minutes seems like poor performance. Another auditory implementation of edge/node graphs is

Metatla *et al.* (2008) where they provide non-visual access to entity relation diagrams, however there is no overview incorporated in the interface and all interaction occurs at the detail level. Brown *et al.* (2006) developed an overview of edge/node graphs where the stated aim was to provide *auditory glances*. The graphs used in evaluating the overview algorithm mostly had eight nodes (more complex than those used in Cohen *et al.* (2006)). The sonifications were generally two seconds long: the left most node was sonified first, then any edges connected to that node, then the nodes connected to those edge, etc. The results showed that participants could correctly pick out a visual counterpart of the graph and that the overview could be used to varying degrees of success for identifying highlighted nodes in the graph. The most problematic graphs were ones with rings. This implies that the sonification algorithm for different topologies needs to be improved. However, Brown *et al.* (2006) showed that a quick representation can provide the gist of the graph without the need for lengthy exploration of each node. Much of the confusion could be alleviated by access to the detailed data. I discuss further the topic of overviews used in isolation versus in conjunction with the detailed data in Sections 3.1.3 and 6.1.1.

Auditory overviews of tabular data

As discussed in the context of spreadsheets (Section 2.1.2), tabular data is difficult to understand using audio. Visually, it is a tidy way of presenting data along two axes – though charts and visualisations often support their interpretation. Kildal (2008) looked at how to provide overviews of tabular numeric data while my work in Nickerson & Hermann (2008) looked at overviews for grid-based games.

In displaying numerical data, there are many mathematical ways to make sense of the data. The challenge is in developing an approach that works generically. Kildal (2008) addresses this by (a) providing a very simple sonification of the data and (b) using an interaction device that supports exploration. TableVis[28] (Kildal & Brewster, 2006a) maps values to pitches; these values are then played quasi-concurrently and give the user a sense of the magnitude of the data. The separation between the notes is minimal resulting in what the user perceives as a single sound, though the user can slow this down through the use of a rotary knob (Kildal & Brewster, 2005). This sonification technique was called High-Density Sonification (HDS). The user can then explore the sounds either column- or row-wise. It is exploration that exposes patterns and trends in the data. The user controls the presentation with a stylus on a touch screen device. This eliminates the need for the audio to express the locations of the rows/columns. This also means that the user can control the speed of the presentation and can go back

---

[28]www.multivis.org

to cross-compare areas of interest.

Kildal (2008) showed that HDS was efficient, effective and could be used with any size data table. Interaction was faster and decreased the mental workload for the users when compared to speech. The usability of TableVis came from a combination of the sonification technique (HDS) and the interaction devices (tablet and stylus). The latter played a big role in the usability of TableVis, as can be seen through the exploratory strategies that participants used (Kildal & Brewster, 2006a) and the guidelines presented in Kildal (2008). The edges of the tablet allowed the user to locate data easily in reference to the limits of the tablet; the ability to draw the stylus across the data gave the user the flexibility to explore at will as opposed to serially. However, the limitations of the system were not fully tested. The questions asked of the participants involved finding where there highest and lowest values were, i.e. where the extremes were. As a result, we do not know if the results hold when interpreting or searching for non-outlier type data.

In Nickerson & Hermann (2008), I used MBS (see Section 2.2.1) to sonify grid-based games. The two games in question were Sudoku[29] and Connect Four[30]. Both games used a tablet and stylus for interaction. In Sudoku, the grid was a simplified 4×4 grid instead of the traditional 9×9 one. Values were mapped to frequency and interacting with a cell excited the grid so that players could hear neighbouring cells. This version of auditory sudoku was informally tested and while the sonification led to a greater sense of the grid, it was not useful enough to constitute an overview. Connect Four, on the other hand, proved playable with only the overview (see Appendix F). The grid was represented as a looping sound where each of the seven columns was a note in a bar of music. The 8th note was silent and indicated the end of the grid. The *x*-axis of the grid was mapped to frequency and each player had her/his own token sound, panned to a different speaker. As the sonification looped through the columns, the tokens in that column were played concurrently and the model used in the sonification helped identify like tokens in neighbouring cells (vertically, horizontally and diagonally). Unlike Kildal (2008), where the overview was tested in conjunction with the detailed data, auditory Connect Four was a stand-alone overview. Without any access to the details (i.e. the ability to query individual cells), players were able to competitively complete games and over half the moves showed evidence of strategy.

These two approaches to providing overviews to tabular data have a certain amount in common: a simple sonification with (quasi-)concurrent presentation of rows/columns aided by tablet/stylus inter-

---

[29]http://en.wikipedia.org/wiki/Sudoku
[30]http://tiny.cc/connect4

action. While the purposes of the overviews were different, they aided users in accomplishing their tasks: identifying high/low values and locating tokens in a line. The ability to draw out patterns from the sounds enabled users to make decisions about where to explore (Kildal, 2008) and where to place tokens in order to win (Nickerson & Hermann, 2008).

Auditory overviews of source code

Unlike the previously discussed overviews, overviews of program source code are targeted at a specific user group: computer programmers. Finlayson & Mellish (2005) and Berman & Gallagher (2009) both investigated the use of sound in order to facilitate familiarisation with complex code.

Finlayson & Mellish (2005) built their work on Vickers & Alty (2005) who aimed to make bugs in code more easily identified. Finlayson & Mellish's (2005) work takes this one step further and provides an auditory overview of the code, called *AudioView*. Earcons were devised for control structures and the overview iterated through the code, showing the structure and hierarchy. They compared non-speech code segment representations to speech and a hybrid approach. With this approach, speech and the hybrid provided better code comprehension and were less taxing on the user. Finlayson (2006) refined the approach to include a fish-eye view. There are no formal results, however early analysis showed that this view provided a better sense of scope and context.

Berman & Gallagher (2009)'s work has a more architectural approach, examining how code interrelates across a whole system. Earcons and auditory icons were devised to represent *beacons* in the code: e.g. methods and classes. They provided various views or filters such as the an overview of the current file, dependent code or code it depended on. The system was evaluated by several experienced software programmers and architects. Participants showed that they could discern the structure of the code and identify local and non-local entities, giving them an idea of how the software was constructed. The audio overview was designed to be used alongside a visual display and showed promise at being able to summarise off-screen information. This allows programmers access to information without changing their current view.

Finlayson (2006) and Berman & Gallagher (2009) show that auditory overviews need to have a level of complexity in order to be effective. This is particularly apparent in Finlayson (2006) where the iterative depiction of source code does not perform as well as when she employs a paradigm from the field of visualisation. With the use of the fish-eye, the user focuses on the detail but has a sense of context. Likewise, in Berman & Gallagher (2009), the various views of filters permit the user to learn

about the whole while not getting distracted from her/his current focus.

Auditory overviews in web browsing

As discussed in Section 2.1.2 [p. 19], auditory navigation of web pages is often done serially and leads to confusion as to what is salient. JAWS, one of the more popular screen readers, provides basic page attributes and calls it an overview. This "overview" lists properties of the page itself and contains no details on the content: it lists number of links, images, whether there is a form, etc. This can give the user an idea of how complex a page is but not what it is about. It is the latter sort of information that can make overviews an asset. I discuss this in detail in Chapter 3.

Pérez-Quiñones *et al.* (2003) address auditory presentation of web pages by focusing on voice interfaces for telephone and similar voice access. Their work is based on adapting the Info Seeking Mantra (Shneiderman, 1996) for auditory information seeking. They replace overview with *situate*, a command to "[help] the user identify his/her location in the information space" (p. 2). The command is used to answer where the user is in on a page and what options are available. Their work was a proposed taxonomy for future implementation and there was no user evaluation of how the situate command worked in practice. However, their main criteria for the overview were location and user options. This implies details rather than overall information about the page.

In a different approach, Murphy (2007) reached the conclusion, in response to participants' requests, that overviews need to be more informative than a listing of attributes. She developed, in conjunction with other researchers, a multi-modal web browser which combined audio with haptics in order to provide an accessible browser for VI users. She found that having an idea of objects on the page or the spatial layout of the page was not sufficient but that a combination of the two provided a model of the hierarchical structure and content that was more meaningful.

As with source code overviews, Murphy (2007) learnt that overviews need to have a certain depth in order to be useful. Most research takes the approach that some meta-data is sufficient. This is reinforced by the vagueness in which overviews are described. Shneiderman (1996) set a precedent by making a circular description of what an overview provides, which is continued in Pérez-Quiñones *et al.* (2003) where the situate command is fuzzily described. This lack of clarity is the main motivation behind my work in Chapter 3 where I formalise the role that overviews play and the characteristics that make overviews useful.

2.3 Summary

Devices and computer interaction paradigms are evolving; there is a need to support devices with smaller or non-existent screens as well as eyes-free interaction for computing on the go. Audio is one modality that can be used to support these types of interactions and is the focus of this thesis due to its ubiquity, ease of generation and the low cost of equipment.

One of the current uses of auditory interfaces is in support of VI users where they have limited or no access to graphical displays. Section 2.1.2 summarises the many ways that auditory interfaces support – and fall short of supporting – VI users. Solving the usability and accessibility problems for this user group can guide us in how to create better auditory interfaces for the general populace.

One research area that emerges from the usability problems with auditory interfaces is getting lost in the details and having to create a mental model of the interface. This is the type of information that graphical interfaces provide via the persistence of visuals. In Section 2.1.3, I proposed that overviews could rectify uncertainty and reduce complexity for the user. This is based on user requests in the research and prior work on overviews in the field of information seeking.

In Section 2.2.2, I reviewed some of the current application areas for auditory interfaces, pointing out where overviews could benefit them. In most cases, the interfaces are usable once the user is familiar with them. However, as discussed in Section 2.1.2 [p. 19], often an unfamiliar interface presents a huge learning curve to VI users, discouraging them from trying out new applications. The main benefit that auditory overviews can provide is moving the user from being mired in the detail to conceptualising the application or interface as a whole. The auditory overviews examined here are varied and while many of the implementations are novel and innovative, confusion comes from the role they play in the interface. It is this confusion and vagueness that I seek to resolve in this thesis: what overviews are (Chapter 3) and how they affect interaction in an auditory interface (Chapter 4).

Chapter 3

What is an overview?

The aim of this thesis is to explore the role that an overview can play in interfaces, particularly auditory ones. In the previous chapter, I mentioned several strands of research that deal with overviews and auditory overviews. Missing from this literature is what exactly is an overview. The term overview is often used without clear understanding of what it means. It is difficult to argue the case for using overviews if it is unclear what they are and what role they play. Working from this premise, this chapter aims to study existing overviews in depth in order to clarify what they should do: their characteristics and function. The literature addressed here and the following analysis covers all types of overviews. By looking at a wide range, the commonalities emerge and it makes it possible to abstract the properties of overviews. This abstract view can then be applied as I have in Chapter 4 where I study how overviews impact the usage of an interface.

To maximise the usage and effectiveness of overviews, it is important to understand what makes them unique. Without clarity on what the term means, anyone can declare an interface an overview and the term becomes watered down. While the term itself is poorly understood, many types of overviews have precise forms. Tables of contents, maps and receipts are examples of overviews that need little explanation: there is a clear expectation of the form of these overviews. Yet, these overviews are rarely called overviews. What follows is an analysis of successful and widely used mainstream overviews – and some less successful counterparts – with the purpose of using generalities about overviews to better understand them and guide in their definition and development.

Webster's Dictionary (Weber, 1984) defines an overview as "a brief comprehensive survey: a sum-

mary" [p. 503]. This definition is vague and can encompass many things, which seems appropriate as the form of an overview can depend greatly on its subject. For example, a table of contents is quite different from a receipt which summarises purchases. However, this definition fails to address the nature of overviews.

In this chapter, I analyse visual overviews by addressing what the literature includes and omits about the nature of overviews (Section 3.1.1) and analyse several overviews in order to better understand what is involved in overviews (Section 3.1.2). The overviews examined are: tables of contents, file managers, line graphs and abstracts. I selected these particular overviews because of their diverse nature: two are graphical representations of detailed information and two are textual. They are also commonplace allowing for an analysis uncomplicated by explaining the purpose of the overview or the detailed information. All the overviews are automatically generated except abstracts, which are handwritten. Each of these overviews is looked at in terms of what it represents, its purpose, its strengths and its weaknesses. I use descriptions of sample usage to show how the overview engages the user. While this list of overviews is not exhaustive, the sample setting is broad enough to understand the various characteristics of overviews. I show that there are six characteristics that span most, if not all, of the overviews described here. This analysis can then be applied to the literature reviewed in Section 3.1.1 in order to determine what is and is not an overview. The second part of the chapter (Section 3.2) then turns to auditory overviews and examines how the six characteristics adapt to a change in modality. I look at how the discovered characteristics apply to the descriptions of auditory overviews in the literature. Each of the auditory overviews shows a majority of the characteristics and their applicability appear relatively unaffected by the change in modality.

3.1   Learning from visual overviews

The majority of established overviews are visual and thus play a great part in informing on the role and characteristics of an overview in general. As such, my search for what an overview is begins with visual overviews. To provide sufficient context for my analysis, I first briefly review the current literature on overviews. This review of current research forms my initial understanding of overviews and guides the analysis of the four common overviews that follows: tables of contents, file managers, line graphs and abstracts.

3.1.1   What we can infer from the literature

As mentioned above, researchers in the field rarely define overviews. The general practice is to use overviews as though it were clearly understood what they engender. However, it is possible to infer the characteristics of an overview as perceived by the authors. For the most part, research involving overviews falls under information-seeking.

Overviews and information-seeking

Shneiderman (Chimera & Shneiderman, 1994; Kules *et al.*, 2006; Kustanowitz & Shneiderman, 2005; Shneiderman, 1994, 1996; Smith *et al.*, 2009) and Plaisant (Hornbæk *et al.*, 2002, 2003; Plaisant *et al.*, 1998b) have focused much of their work on information-seeking and visualisations of large data sets. Jointly, they have published many papers on the necessity of facilitating searches in large collections (**North** *et al.*, 1996; Plaisant & Shneiderman, 1995; Plaisant *et al.*, 1996, 1998a). They recognise that overviews are important when performing information seeking tasks. Shneiderman in particular champions overviews as an integral part of the interface. However, overviews have never been their main focus. Many of their publications state that overviews should be included and that different types of overviews play different roles.

Overviews are put into prominence in Shneiderman's 1996 paper about the information-seeking mantra (**Shneiderman**, 1996). They are highlighted as the first step in interaction: "overview first, zoom and filter, and then details-on-demand" [p. 1]. The paper only states that overviews should be present, not what form they should take nor what information they convey. Instead, Shneiderman explains that "overviews [allow users to] gain an overview of the entire collection" [p. 2] This statement makes it clear that overviews are comprehensive and span the entire data set rather than a subset. Meanwhile, in another paper, he and his fellow authors go into more detail when stating that an overview "provides context"[p. 79], "provides general understanding of the contents" [p. 76], "guides [the user] to desired location" [p. 79] and "promotes exploration" [p. 79] (**North** *et al.*, 1996). The basis of these assertions is not revealed, but many of these attributes are confirmed by my analysis later in this chapter.

Subsequent papers deal with furthering information visualisation, with a focus on developing ways to bring out relationships within collections. For example, **Plaisant** *et al.* (1996) describes a system for overviewing personal histories using a time scale. The timeline allows practitioners to relate disparate events and build a more complete picture of a person's medical or other history. The focus of the interface is on completeness, while reducing complexity through filters. Another aspect of the interface

is to show importance so that salient records are easier to find. Additionally, patterns in the data emerge when interaction with a record highlights related records.

In Hornbæk *et al.* (2002), the authors summarise guidelines for overviews. However, these guidelines refer to look and feel of overviews rather than what should be included in one. For example, one guideline discusses the size of the overview window in the interface so not as to limit the amount of the detailed information that can be seen. Such guidelines are useful in the design of interfaces but do not provide insight into overviews themselves.

There are various themes that run through the work that Shneiderman, Plaisant and their colleagues have done on information seeking. The first is that overviews are a key aspect of an interface, visual and auditory alike. Second is that they are comprehensive and describe the entire data set. Third, they promote exploration by showing interrelationships. And finally, they play a guiding role, facilitating navigation. Though they do not state it explicitly, overviews are also brief.

Thumbnails as overviews

There are other strands of research that touch on overviews. One common use of overviews is for exploration of a 2D graphic, such as a map. In these cases, there will often be a thumbnail of the entire graphic in one corner of the interface. The user explores a zoomed in subset of the graphic, as shown in Figure 3.1. For example, Hornbæk & Frøkjær (2001) describe a comparison between three interfaces (detail only, fish-eye and overview+detail) for the reading of electronic documents. Here, the overview is not a table of contents or other such abstraction but rather a thumbnail of the entire document where headers stand out due to their larger size. In this case, overview seems to be synonymous with miniaturised. This is also the case in an ACM Computing Survey where (Cockburn, Karlson & Bederson, 2008) provide a comparison between overview+detail, zooming, and focus+context interfaces. They use mapping applications with overviews as their primary example of overview+detail interfaces. What the paper makes clear is that overviews are separate from the detail, unlike zooming and focus+context where the user interacts directly with a single view.

What sets this work apart from the other overview work described in this section is the fact that the overview, or more appropriately the thumbnail, does not provide different or re-organised information. The thumbnail is a scaled down version of the whole which enables the user to see the entirety of the detailed data at the same time. However, it does little else. It does not summarise or provide an abstracted view of the detail. Thumbnails are simplistic in a way that the other overviews are not. Based on my

Figure 3.1: **A map with thumbnail** Source: http://globalr2p.org/ ©2009-2010 Ralph Bunche Institute for International Studies

analysis later in this chapter, I would argue that these are not overviews. Unlike Plaisant's overview of personal histories (Plaisant *et al.*, 1996) , there is little richness provided here. This is addressed further at the end of Section 3.1.3.

Attributes of overviews from the literature

The literature reviewed above deals with overviews however it uses them without clearly delineating what they are. The purpose of this review is to draw out what can be generalised about overviews. There are several themes that crop up in most of the research: first is that overviews are important in navigating the data set; second is that most research refers to the overview as being comprehensive or representative of the whole; and third is that overviews show salient features and interrelationships. From these three points, an impression of the role and characteristics of overviews emerges.

The importance of overviews   All the research refers in some way to the overview being important to the usability or usefulness of the interface. The overview plays a guiding role, making the content more comprehensible. It also provides access to the detailed information by showing its salient features.

Representing the entire collection   That overviews are to be comprehensive is also quite universal to the research reviewed here. Shneiderman (1996) and Plaisant *et al.* (1996) state it most explicitly, talking

about the entire collection and focusing on completeness. In the rest of the research, representing the entire dataset remains an unstated theme.

Salient features and interrelationships   One of the other re-occurring themes in the literature is that overviews expose interrelationships and key features from the detailed data. Plaisant *et al.* (1996) in particular put an emphasis on showing the importance of relationships between events in a personal history so that salient records emerge. This is also the case in North *et al.* (1996) where the overview is a guide to a collection of anatomic images.

3.1.2   An analysis of common visual overviews

The previous section summarises much of the current research into overviews. While researchers generally agree that overviews are important and describe some of the advantages of an overview, few address overviews directly; they describe the effect of the overviews they developed. This section delves into overviews as a concept. What is an overview? What characteristics does it have? What role does it play in interaction? Given that auditory overviews are not firmly established, I turn to their visual counterparts for examples and this section confronts these questions by looking at four common overviews. These overviews are well-known, need little introduction and are well-understood. The first of these overviews is tables of contents which are a text based overview of organised written material; next, I look at file managers which are an interactive representation of computer file systems; third are line graphs which are a graphical representation of data; finally, I focus on abstracts which are hand-written textual descriptions. I analyse typical interaction for each, identify their components and representation and discuss what each implies about overviews in general. Section 3.2 then returns to the auditory modality to examine if the findings are universal and not specific to visual overviews.

Tables of contents

Tables of contents have long been considered to be a vital component of written material in order to allow the reader to peruse it. They are an example of a well-known overview that is understood to such a degree that they can be automatically generated. This is aided by the fact that the detailed information is very structured: the headers can be extracted and displayed using the components described in Figure 3.2. While context of use drives the interaction with a table of contents, the basics of glancing at headers and absorbing the structure remain. The speed of interaction and attention to detail will vary. What follows is a typical interaction with a book.

headings and subheadings · sets of headers · heading numbers · page numbers · leaders · indentation

Figure 3.2: **A sample table of contents** showing the various components described in table 3.1

Typical interaction   A reader picks up a book and turns to the table of contents. S/he peruses the text, letting words of interest capture her/his eye. When s/he notices something interesting (whether it be by chance or by directed search), s/he takes a closer look. The title of the chapter or section gives an idea of how the topic is handled. Its position in the hierarchy of the book determines how significant the topic is in the scope of the book. Additionally, the page numbers indicate how many pages the discussion of the topic spans. If sufficiently interested, s/he may flip to the pages in question and either read or scan the text. Typical questions are *what (else) is this book about?* and *where is there information about <topic>?*

Components and their representation   The amount of information displayed in a table of contents is driven by the type of book. A piece of fiction may have no table of contents at all, or possibly an intentionally cryptic one. In non-fiction, the table of contents is usually thorough and explicit.

In general, a table of contents will be made up of broad information about the subject matter of the book: preliminary matter such as the foreward and lists of illustrations/figures; the titles of parts, chapters and sections; end matter such as the bibliography, indices and end notes. Each item can be accompanied by its page number and section number. In *Copy-editing: The Cambridge Handbook for Editors, Authors and Publishers* (1992), Butcher specifies that the table of contents should be comprehensive as "it is easier for the reader to have only one place to look" [p. 177] but warns against the list being too cluttered and general practice should be to limit to only one level of subheadings. If the book has multiple authors, Butcher states that the table of contents may include the authors' names and possibly titles. Headings and subheadings can be emphasised or differentiated by variation in font size or weight. Level of indentation and number schemes also describe their place in the hierarchy. To ease reading, there are often *leaders*; these are normally in the form of periods that run from the section title

Component	Role
Headings and sub-headings	shows topics covered
Sets of headers	flow of topics
Heading numbers	structure of the material
Page numbers	where material is located
	also depth in which material is covered
Indentation	differentiates major topics from minor topics
Leaders	legibility: which pages numbers correspond to what topic

Table 3.1: **Components of a table of contents and their roles in the overview.** These are shown in Figure 3.2.

to the page number. In Table 3.1, I list the various components of a table of contents and their function as part of an overview.

Discussion   Superficially, tables of contents look dry and simple: they are listings of the top level headers of a document. One might assume that tables of contents are merely an enumeration. However, when writing, an author or editor may ask "is this an adequate representation of the work?" The answer to this question would be reflected in changing the header names, not the table of contents itself. Therefore, one aspect of the table of contents is how it reflects the work. When one looks more closely at how to interact with the table of contents, it is noticeable how it enables exploration and provides fresh viewpoints where the reader discovers things that s/he did not know s/he was looking for (similar findings are described in Kwasnik (1992)). Much like the headers of each page of a dictionary, eyes get distracted and caught by other topics that might or might not be related to the information sought. In both instances, the overview encapsulates what the document is about and acts as a guide. One of the reasons we find tables of contents useful is how they allow us to quickly place each section in context and explore what dances around the topic in a way that an index does not. Contrasting a physical table of contents to a web one highlights the effect of page numbers on the interaction. On paper, there are defined page numbers: this can communicate size; online, there is no indication how much information will be contained in each section. This is where a static paper document can provide a more informative table of contents than a dynamic, digitised document. In essence, the flexibility of the digitised material breaks one of the strengths of tables of contents. Though, I should clarify that the digitised version can offer other advantages, such as being able to navigate directly to a page of interest in the document.

So how is an overview different from an enumeration? A table of contents does not simply list headers, but also describes the structure and nature of the document. It shows context as well, and by showing this context, it can pique the reader's curiosity about other topics. This is a strength of an

organised overview rather than a search feature that simply presents an answer. This can be seen more clearly when examining works of fiction. Textbooks and non-fiction usually have a table of contents. This is not necessarily the case when speaking of fiction. A collection of stories or poetry might have one but a novel might not. Why? One reason is that there might be no chapter titles. If there are chapter titles, it begs the question of why a table of contents was not considered important. In other types of books, the title of the chapter provides no information yet is listed in the table of contents. I argue that this can be explained through the concept of discovery and exploration: the surprise of the chapter title is an important part of the experience and including a table of contents might give too much away. In cases where there is a table of contents, it sets the mood of the book or intrigues the reader by providing hints as to what is to come. This concept is discussed in (Barchas, 2003), where the author describes how the table of contents adds to the reader experience.

Tables of contents also enable a reader to find where s/he left off in a book. The chapter/section titles identify where to pick up reading again. In this case, the table of contents contains the reading history. A reader might scan the table of contents rejecting certain sections as already read or not read yet with the goal of pinning down where s/he left off reading.

Based on this analysis of tables of contents, I argue that overviews include the following qualities: describing the scope of the material, describing the structure of the material, describing the hierarchy the material (if one is present), quickly guiding the reader to topics of interest, showing contextual information about topics, helping the reader find where s/he left off in prior interaction and encouraging the reader to explore other topics.

File managers

The advent of graphical user interfaces brought a change in the representation of the computer file system. The file manager in the Windows[1] operating system displays the current path in the address bar, the contents of that directory in the main panel and an interactive representation of the directory tree structure on the left. This representation of the directory tree structure is an overview that describes the entire contents of the file system where the user can choose which portions to hide or expand depending on her/his task (see Figure 3.3). The main advantage of the overview is the visibility of the entire file system where previously the user would have to stitch together her/his own internal representation of the surrounding context.

---

[1] http://www.microsoft.com
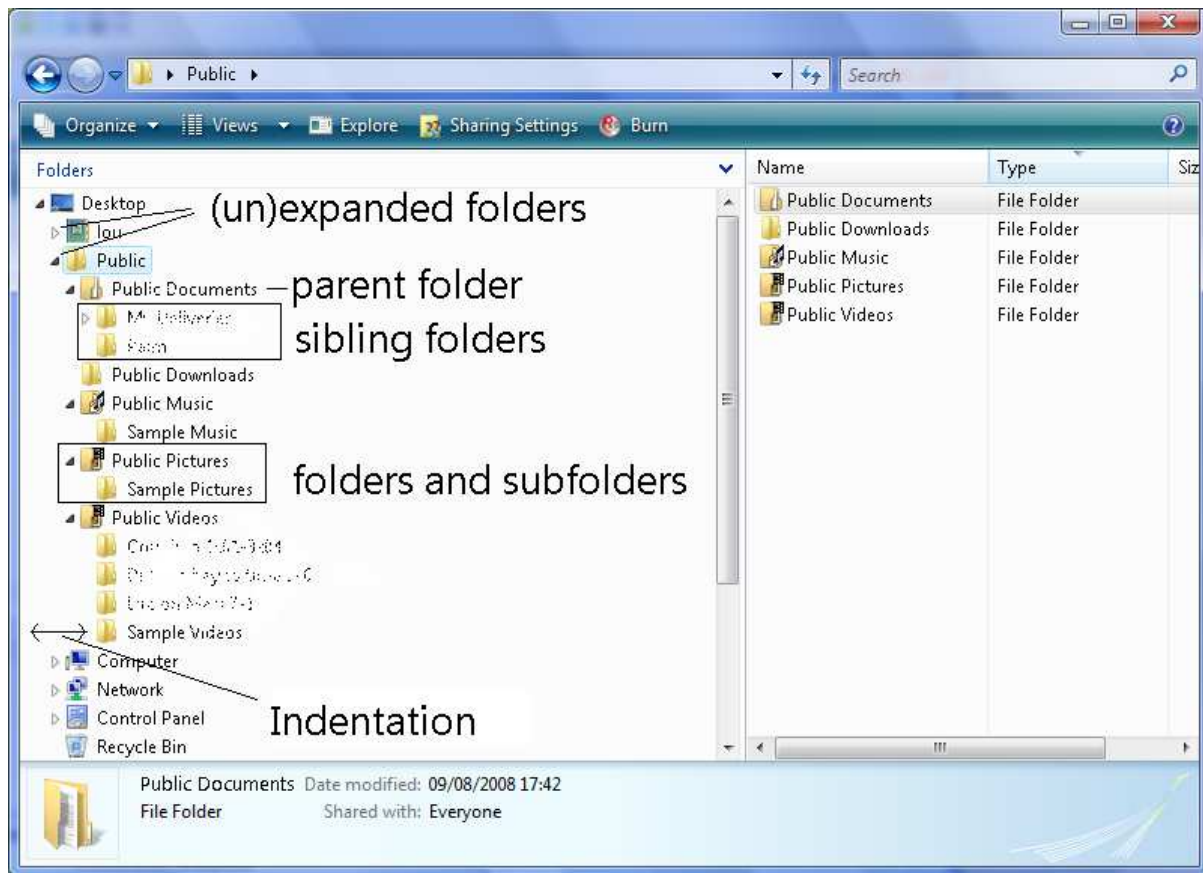
Figure 3.3: **A file manager window** showing the various components described in Table 3.2

Typical interaction   A typical task one performs in the file manager is locating a file. Upon launching the file manager, a user might use the directory tree overview to locate the part of the computer file system to be explored. Default views vary from showing the root of the directory tree to showing the home directory. Regardless of the initial view, the user would scroll and hide/expand folders of interest. Assuming the folder of interest is several levels deep, s/he would expand the parent directory to see the folder of interest and its siblings. S/he would then be aware how deep the file is in the directory tree structure. The name of the parent directory and the sibling folders give an indication of the organisation surrounding the folder of interest. Then, s/he could click on the folder to show its contents in the main panel and navigate to the file of interest.

Components and their representation   In general, file managers and their overviews of the directory tree structure are simple and straightforward. Their strength comes from the way they display the relationships between folders so that users are familiar with the organisation and have an easier time remembering or deducing where they stored certain files. Research into recognition and recall shows

Component	Role
Folders and sub-folders	shows folders in the directory tree
Sets of parent and children folders	organisation of information
Sibling folders	folders loosely associated with each other
Parent folders	general category of a folder
Expanded icon	state of the folder (expanded or hidden)
Indentation	folder depth in the directory tree

Table 3.2: **Components of a file manager and their roles in the overview.** These are shown in Figure 3.3.

that a prompt can bring to mind information that otherwise is considered forgotten (Tulving & Pearlstone, 1966). Broadly, the main components of the directory tree overview in the file manager are the icons that represent directories and whether they are expanded or not. The organisation of the directory tree is mostly made apparent through the spatial layout that indicates the relationships between folders. These components are described in Table 3.2.

The most obvious aspects of the overview are the folders and sub-folders.Through the parent or sibling relationship one can get an impression of the logic behind the organisation and naming scheme. An icon next to the folder name, in the shape of a triangle or a plus or minus sign, indicates that the folder has children. A downward pointing triangle or minus sign indicates that the folder is expanded and the children folders are displayed; a triangle pointing to the right or a plus sign indicates that the contents of the folder are hidden. The relationship between folders is further reinforced by indentation. A sub-folder is indented to indicate that it belongs to the parent directory.

Discussion   The directory tree overview is a relatively new overview when compared with tables of contents and there have been many criticisms of the file manager overview in the academic community (e.g. Barreau & Nardi 1995; Freeman & Gelernter 1996; Hicks *et al.* 2008). The metaphor of the file manager is based on the real-world organisation of office work. The file manager is a filing cabinet, directories are folders and the contents are files. The novelty is the digital representation of these real world artefacts. The main criticism is the difficulty of finding files quickly and efficiently. This is an action that the overview should support.

One flaw is a break in the hierarchical organisation: the Windows file manager describes the Desktop folder as the top level directory when in fact the Desktop folder is a child of the user's home directory. As a result, the Desktop folder appears twice: at the top level and as a third-level directory, e.g. `/Users/<username>/Desktop`. The overview is meant to give users a way of conceptualising

the whole file system and presumably the developers of the file manager wanted to reinforce using the Desktop as a point of departure. However, this breaks the actual hierarchy of the directory tree. It has the potential effect to cause confusion. How can a directory be its own great-great-grandparent? This is handled differently in OSX[2], where the main overview is divided into sections (see Figure 3.4. This allows the user to choose to navigate from the top level – `Macintosh HD` – or from places – the standard ones being `Desktop`, `Home`, `Documents` and `Applications`. The addition of key places removes the necessity to break the model yet allows for quick access to places of interest. However, as I discuss below, this view does not describe the file system fully which seems to be an important part of an overview.

Additionally, as sections of the file system are expanded, portions of the directory tree structure disappear off the edges of the overview pane. At that point, the user must scroll up and down or left and right to see the full context of a directory. On the surface, there appears to be no difference between scrolling on the screen and flipping between pages in a paper table of contents. However, the difference is that items change place on the screen while information on paper is static. This can result in confusion until the user finds her/his place again in the view. The scrolling from left to right can also be disconcerting. This occurs when folder names are very long or the user has expanded folders quite deep in the file system. The user can then adjust the width of the overview pane or scroll to see more of the content. However, this is more of an implementation usability problem – where there is a trade-off in usage of available screen space – than a problem with the overview per se.

As mentioned above, the Apple file manager takes a different approach (see Figure 3.4). The default view uses a series of panes where each pane lists a level in the directory tree. The result is that the user can see several levels back from the current directory being explored but the view is limited to the context of the current path. For example, in a directory that is three levels deep, the user would see the contents of the top level directory, the contents of the parent directory and the directory in question. This approach has advantages and disadvantages but in terms of overviews it does not qualify as it acts as a filter as opposed to an overall representation of the file system. More specifically, it acts more as a focus+context interface, where the user sees the current detail in full focus but also gets contextual information. (Usability First, 2011). The distinction is that the context provides a view of surrounding information but not of the whole.

So what does the Windows file manager imply about overviews? The analysis of tables of contents

---

[2]http://www.apple.com/macosx/
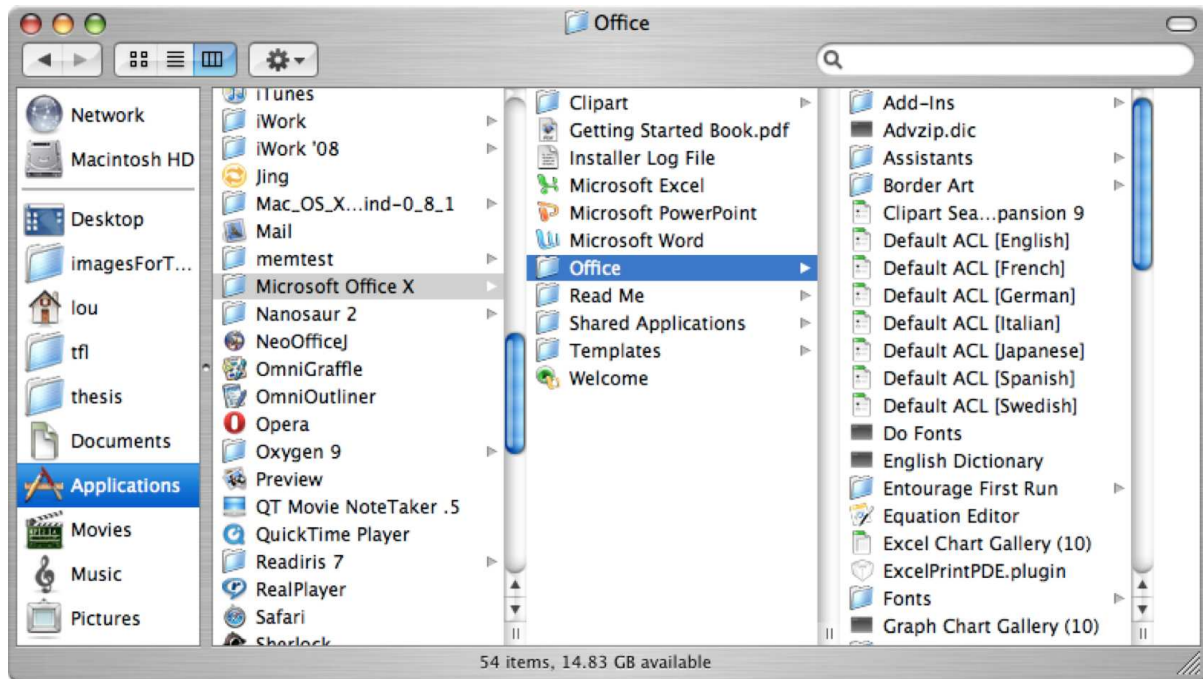
Figure 3.4: **A finder window** showing 3 panes of directories. The overall organisation of the file system is not visible.

revealed that an overview describes scope, structure and hierarchy. The file manager does these as well though the interaction degenerates when parts of the overview scroll out of its pane. Building on the analysis of tables of contents, the file system overview reinforces the descriptive aspects of an overview but does not support or refute the exploratory aspects. However, it does bring out a certain historical aspect that is harder to understand when dealing with a static system. While not immediately obvious, tables of contents also have this aspect of providing a historical snapshot. Where there is a version control system, the history of the document's creation is described by the evolution of the table of contents. One reason it is less obvious is that a table of contents at some point achieves a final state while a file system is generally always in flux. The file manager overview describes a living and changing dynamic system that is largely organised by the user. As a result, when not shared with others, it is less critical that it be clear and flow well. Whereas with a table of contents, the flow and sharing is of vital importance from creation.

Line graphs

Line graphs, such as seen in Figure 3.5, are a visualisation tool for showing interrelations between two measures. By plotting data against the two axes (x and y), connections between the two measures often become more apparent and can strengthen the presentation of findings. Generally, line graphs are
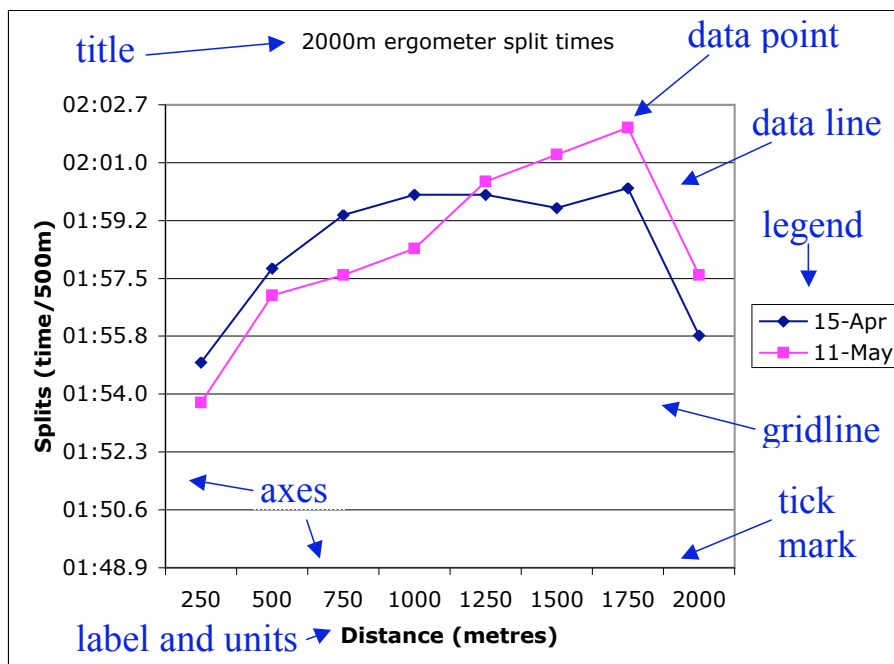
Figure 3.5: **A sample line graph** showing the various components described in Table 3.3

considered to be a visualisation however they are also an overview of a data set. Line graphs, which are both overviews and visualisations, are made up of complex components but are displayed very simply and intuitively. Like tables of contents, they are a well understood overview that have been in use for a long time. Their usage in statistics and analysis made them a key feature in spreadsheet programs such as Excel[3] and they are well enough understood to be able to be generated automatically.

Typical interaction   When a person first encounters a graph, it is generally as an introduction to some information or to support some text. S/he would glance at the line taking in its peaks and valleys and interpreting its meaning. As s/he studies the graph further, s/he may jump from glancing to studying the various components of the graph. S/he may want to understand the relationship between what each axis represents, and if so s/he would be driven by questions such as "how does <x> affect <y>?" As questions get more specific, her/his focus shifts to subsections of the graph. Lohse (1991) details the cognitive processes used in graph interpretation, discussing how one's focus shifts depending on the question.

Components and their representation   A simple line graph contains more elements than is immediately obvious. The main component is the line which is made up of data points and the lines that connect them. The positions of the data points are determined by the scales of the two axes. The *x*-axis usually

---

[3]http://office.microsoft.com/en-us/excel/

Component	Role
Line	connect the data points and show the progression or variation of the data along the two axes
Data points	show the specific values for each measurement
Axes	show what is being compared
Axis labels and units	describe the axes and their scale
Grid lines/tick marks	show the scale of the axes and help people to approximate data point values
Title	describe the subject of the graph
Legend	labels the data line in the graph

Table 3.3: **Components of a line graph and their roles in the overview.** These are shown in Figure 3.5.

runs from left to right and the *y*-axis runs from bottom to top. The axes will have labels and units of measurements so that one can tell what the data points mean. To make the graph more legible, the axes are generally marked up with grid lines or tick marks corresponding to the units of measurement which help viewers approximate the values that constitute each data point. A title and a legend further describe what the line graph is showing. The title is a description of the graph and the legend specifically describes the line in the graph.

A more complex line graph might include multiple lines as well as advanced components such as trendlines, which further analyse the data, or error bars, which show deviations in the data. These variations of the basic theme serve to help the viewer to pick out details of interest. Table 3.3 shows the main components – but not these more advanced ones – of a line graph.

Discussion   In terms of the previously exposed attributes of overviews, line graphs describe the scope and structure of the data. In addition, they quickly guide and entice the viewer and provide a snapshot of the detailed information. However, they sometimes only describe a subsection of an entire data set. When the data set is small, a graph can describe it in its entirety but as the data increases in dimensions, it may only describe one aspect.

Regardless of the scope of the line graph, it can depict a multitude of information. Like a table of contents, a line graph invites exploration. Initial interaction can be cursory: the viewer glances at it to see what story it has to tell. As interaction deepens, the viewer can begin to study the graph for answers to new questions. Graphs are often thought of as a way to show findings in the data; however, graphs can be equally powerful at identifying areas of interest to analyse. In other words, graphs play a role at the beginning of analysis by showing the researcher where to look as well as at the end of research as a way to convince an audience of some finding. Cairns & Cox (2008) discuss graphs as

tools in exploratory data analysis, specifically highlighting their ability to provide "a good view of the relationships and oddities in the data from experiments" [p. 120]. Graphs can also play a role in error checking calculations: should some statistic be miscalculated, graphing the data can expose the error by not taking the expected shape. This is an aspect of overviews that I described in relation to tables of contents: their ability to expose aspects of the detailed information that are not immediately obvious without the benefit of an overview.

Abstracts

Abstracts are textual, like tables of contents, yet are hand-generated. This is in contrast to the previous examples which can all be generated automatically from the detailed information. They are generally short texts of 200 to 300 words that describe a piece of work, for example an article or a museum exhibit. They are carefully constructed to provide a good representation of the work in question and serve as an enticement for people to explore the full work. For example, an abstract of an article encourages the reader to read the full text. In fact, most publishers, such as the IEEE[4], only allow access to the full article until a fee has been paid and they use the abstracts in part to convince people to read or buy articles.

Typical interaction There are many ways to write an abstract. This usage is derived from multiple guidelines on writing abstracts (Colorado State University, 2011; Koopman, 2011; University of North Carolina, 2011). When presented with an abstract, a reader will scan it for interesting keywords. If interested, s/he will read the abstract in more depth. Depending on the subject matter, it might be necessary to re-read it several times to gain a good understanding of the content. By glancing at the opening sentence, s/he gets an idea of whether the subject matter is of interest. With the first read, the reader gets a general grasp of the topic at hand. On subsequent reads, s/he gets a better grasp of the work by studying the message of the abstract and reviewing the central explanatory part of the abstract.

Components and their representation Abstracts are very specialised and good ones are carefully constructed to best represent the work in question. Despite their need to be highly individualised, there are certain generalisations one can make about them. In the most general terms, I argue that an abstract will have an introductory statement, a main message and an explanatory section. The order and format depends greatly on the type of work, the forum/audience and the perspective of the author.

   The introductory statement describes the purpose of the work and sets the scene. It is intended to

---

[4]http://www.computer.org/portal/web/csdl/home

Component	Role
Introductory statement	describes the purpose of the work and sets the scene
Main message	what the author wishes the reader to remember
Explanatory section	describes background information and/or methodology

Table 3.4: **Components of an abstract and their roles in the overview.**

grab the reader's attention. The main message is what the author wishes the reader to remember. It may be a research finding, a new revelation or the contribution of the work. The explanatory section describes background information and/or methodology.

Discussion Because of the prose format, abstracts are harder to create and harder to describe than the other overviews studied in this chapter. The format depends on the subject matter but they are always short and the author must be very concise. The one exception is extended abstracts, which are more properly introductory papers rather than an overview. Unlike the other overviews in this chapter, abstracts resist automatic generation. With the other overviews, there is a simple mapping of detailed information to their abstraction. Automatically generating abstracts or summaries is a rich area of research in Information Retrieval but as of yet, no computer-generated abstract is reliably as powerful as a hand-generated one (Jones, 2007). As such, abstracts are a good example of how one needs understanding of the underlying work to be able to generate the overview. With numerical data, the fields of mathematics and statistics provide a way to interpret it. With tables of contents and file managers, the overviews stem from the self-imposed organisation of the data: section headers and directory names. With prose, there is no straightforward way to summarise the text. The way we express ourselves in words is not yet well enough understood to negate the need for a custom overview. Similarly, in cases where we are aiming to form new overviews, overviews will need to be custom generated until the underlying format is well enough understood to be automatically processed.

3.1.3 Results of the analysis

Each of the overviews described in this chapter provided insight into the nature of overviews and what they do. Table 3.5 is a summary of all the major characteristics that emerged and their presence in each overview. It is apparent that some characteristics are more common than others. There are three characteristics that appear in all the overviews and three that appear in half of the overviews. The most common, i.e. those that appear in all the overviews I studied, are 1) showing scope, 2) quickly guiding and 3) showing context. These are core to building an appropriate overview. The characteristics that

Attribute/characteristic	General category	Tables of contents	File systems	Line graphs	Abstracts	
Scope of the material	*descriptive*	x	x	x	x	100%
Quickly guides to information	*exploratory*	x	x	x	x	100%
Shows contextual information	*exploratory*	x	x	x	x	100%
Exposing the structure of the material	*descriptive*	x	x	x	_	75%
Encourages exploration of other information	*exploratory*	x	_	x	x	75%
Provides a snapshot of the state at a particular time	*historical*	x	x	x	_	75%

Table 3.5: **The attributes and characteristics of an overview based on the analysis in this chapter.** The final column shows how often the attribute is represented in the overviews surveyed.

do not appear all the time are less fundamental but still important. They are 4) describing the structure, 5) encouraging exploration and 6) providing a historical snapshot. What follows is a discussion of each of these main characteristics. These characteristics can then guide overview development by defining what an overview *should* do.

Showing the scope of the material

Attribute/characteristic	General category	Tables of contents	File systems	Line graphs	Abstracts	
Scope of the material	*descriptive*	x	x	x	x	100%

Table 3.6: Showing the scope of the material

Showing scope is about completeness and delimiting the data set. This attribute picks up the theme that emerged from the review of the literature: representing the entire collection. By describing the whole, an overview sets initial expectations and filters can hone in on subsets of the material. This is one of the key points set forth by **Shneiderman** (1996) about overviews and is the common theme amongst all the research surveyed in Section 3.1.1. Showing the scope is about getting to know the whole data set. Further drilling down is part of filtering the overview so that the user can examine an area of interest in greater detail. For example, in tables of contents, a reader can flip through it and her/his eyes will perform filtering by ignoring information that is not of interest. Additionally, with a line graph, a viewer can see the first and last data points. This framing allows the viewer to begin to process what comes between. In an abstract, setting the scope is even more important: it describes what is contained within and will guide readers' expectations of what the article, for example, will and, more pertinently, will not contain.

Acting as a guide

Attribute/characteristic	General category	Tables of contents	File systems	Line graphs	Abstracts	
Quickly guides to information	*exploratory*	x	x	x	x	100%

Table 3.7: Quickly guiding to information.

Quickly guiding is an important role an overview can play. This picks up on two themes from the literature: showing salient features and obscuring detail. The overview should indicate whether there is information of interest in the detailed data and how a user might go about a certain task. One of the main reasons for overviews is the difficulty of comprehending the whole data set; the abstraction that an overview provides as well as its brevity allows a user to quickly see patterns and relationships that would be harder to see if s/he were perusing the detailed data. By presenting higher level information and doing so in a brief manner, the overview quickly guides to the detailed information. For example, in a line graph, the shape of the line allows the viewer to pinpoint more easily trends in the data and to identify areas worthy of further examination. Likewise, a table of contents pinpoints the location of certain information. Only from in depth knowledge of the detailed information could a user glean the sort of understanding that is readily available from an overview.

Showing contextual information

Attribute/characteristic	General category	Tables of contents	File systems	Line graphs	Abstracts	
Shows contextual information	*exploratory*	x	x	x	x	100%

Table 3.8: Showing contextual information

Having a sense of context allows the user to better understand what is being presented and how it is addressed in the data set. If one wanted to know if some information were present, a search feature would be sufficient. At that point, the detailed data is like a black box where a user dips in to find out some information but does not know what else is there. For example, in a table of contents, seeing sibling topics, i.e. headings in the same branch of the tree structure, gives an idea of importance or depth in which a topic is covered. Likewise, with a line graph, one can see at a glance where the majority of data points fall and outliers are immediately identifiable.

Attribute/characteristic	General category	Tables of contents	File systems	Line graphs	Abstracts	
Exposing the structure of the material	*descriptive*	x	x	x	–	75%

Table 3.9: Exposing the structure

Exposing the structure of the material

The structure or organisation of the detailed information should be exposed in an overview. Much as the scope delimits the detailed data, the structure helps the user understand what comes between those limits. While context is about what relates to a particular piece of information, structure is about the flow of the detailed information. The extracted high level information in the overview should reflect this lower level organisation. As with the scope and context, structure manages the user's expectations and allows her/him to make sense of the entire data set.

For example, a user may not want to explore all the folders in a file system but understanding the structural logic behind the organisation gives her/him an idea of the similarities or disparities between content. In a work scenario, project folders might follow a particular pattern containing folders for the different phases of the project. Seeing the organisation means the user can easily situate where a particular folder might be located when exploring a new project.

However, abstracts do not expose the structure the same way as other overviews do. In this case, it would be the role of the introductory section to explain the flow of the document. The prose format of an abstract makes it more difficult to smoothly integrate organisational details into the text. The structure or flow of the document might be hinted at in the abstract, but this is incidental and not a requirement of most abstracts. The rigidity of including structural information could cause the abstract to be formulaic and less representative of the work. In a sense, an abstract plays the dual role of overview and marketing material. In order to do the latter well, an abstract must be free to vary its format, while still containing the main components (see Table 3.4). The order and form may change: it is how well it gets across its message that is important. In other words, while structural information is key, it is secondary to the main purpose of an overview which is to best describe the detailed data.

Encouraging exploration

One of the main purposes of an overview is to act as a guide; encouraging exploration is a corollary to this. The overview should make it possible to discover what is there: not only what the user is looking for, but also what *else* is there. In other words, what surrounds the topic of interest; the contextual

Attribute/characteristic	General category	Tables of contents	File systems	Line graphs	Abstracts	
Encourages exploration of other information	*exploratory*	x	_	x	x	75%

Table 3.10: Encouraging exploration

information implies how the topic is handled. To achieve this goal, the overview should encourage exploration and engage the user's curiosity. For the most part this is easy to do simply by presenting higher level information. Headings in tables of contents draw a user's eye as s/he looks for topics of interest; the descriptive text (such as the legend and labels) and the shape of the line in a graph encourage the viewer to study it further; the story told in the abstract entices the reader to read it again in more detail and to explore the article.

Both tables of contents and abstracts can describe a written document. However, each goes about it in a very different manner. The abstract is prose and describes the text. Its purpose is to capture the nature of the document. On the other hand, the table of contents is more focused on exposing the structure and topics in the document. Its purpose is to allow readers to quickly jump to a section of interest. Both, however, encourage exploration: the abstract by presenting concepts and enticing the reader to discover the story within and a table of contents by letting the reader peruse the subject matter and exposing how subtopics are addressed.

Another facet of encouraging exploration is the difference between a known and unknown dataset. For example, a document author experiences the table of contents differently than a reader new to the document. As an author, I find studying my table of contents improves the flow of my document and impacts decisions in the text. Levine's thesis writing guide (2010) supports this in describing how authors can "improve [their] manuscript" by generating and using tables of contents in their writing. In this sense, header naming plays more of an role in the table of contents than in the document itself. That is to say an improperly named header may only become apparent when seen in the context of the table of contents. Therefore, encouraging exploration is a factor of mode of use: when an outsider, the overview should encourage exploration and when an insider, or creator, the overview will not engage the curiosity in the same way but will instead provide insights that the insider was perhaps not aware of.

Providing historical states

Overviews can help track changes in time. When a data set is dynamic, such as a file system or a working document, an overview can capture the state of the data at a particular time. Tracking changes through

Attribute/characteristic	General category	Tables of contents	File systems	Line graphs	Abstracts	
Provides a snapshot of the state at a particular time	*historical*	x	x	x	–	75%

Table 3.11: Providing a snapshot of the state at a particular time

the detailed data can be cumbersome and the overall meaning of those changes difficult to understand. Whereas an overview provides an easy way to capture the general, if not the specific, changes to the data. For example, in a document such as this thesis, the state of the table of contents seen over time can give an indication of progress. As headings appear and change, the feel of the document changes, exposing areas of activity. This type of historical state can then be useful to expose the evolution of a data set from one version to the next.

Levels of overviews

So far, I have discussed overviews in general. The aim was to understand overviews as a concept. However, once overviews are used in practice, the context can have a great effect and can muddy the waters. For example, a table of contents describes a text and a portion of the table of contents can describe a subsection of that text. Does that mean this portion of the table of contents is not an overview? The six characteristics state that an overview should be comprehensive: comprehensive with respect to the text in question. It therefore follows that there can be sub-overviews.

When is an overview a sub-overview? Logically, a sub-overview provides a comprehensive overview of a subset of a larger dataset. For example, a table of contents could describe a book and a sub-overview might describe a chapter. The sub-overview might take a different form than in the table of contents as the chapter becomes the whole instead of being a subset of the book. In other words, the sub-overview becomes an overview in its own right. Defining it as a sub-overview is only necessary when discussing both the overview of the book and the overview of the chapter at the same time.

Task dependency The discussion of sub-overviews above highlights that different overviews support different tasks. Looking at a paragraph may only be useful in particular context; conversely, understanding an entire book may be useless in another. In other words, the task and the overview need to be closely related. This task dependency is included in the six characteristics. This is most obvious in the case of quickly guiding and encouraging exploration: neither of these can be accomplished well without the overview being appropriate for the task in question.

Most of the overviews analysed here are general and modular. They are applicable to a variety of tasks but are not suited to all tasks involving the detailed data they describe. For example, both tables of contents and abstracts can describe a text but they do not serve the same purpose. If the task is to understand the thesis of a text, a table of contents may provide hints but not as well as an abstract could. A successful general overview does a better job of satisfying all the characteristics if the overview anticipates the tasks for which it will be used. A city map could help in several tasks like route-finding, understanding the layout of a neighbourhood or understanding the network of arteries through and around the city. Graphs, on the other hand, are harder to make universal and are often fine-tuned to a task. Another example is a timeline which shows a sequence of events but not interrelationships between them. However, in the case of the timeline, trying to make it more general might compromise in its quality: highlighting connections between events or related people could obscure the sequence of events. A timeline quickly guides to the *when* but not the *how*. If the task is understanding the former, then the user is more likely to be encouraged to explore.

Satisfying the six characteristics aids in assuring that the overview is suitable and useful. Examining the task and overview together and checking, in particular, for the more interactive characteristics, helps determine if the overview is the correct one for the task.

Why thumbnails are not overviews

Previously, in Section 3.1.1, I discussed thumbnails as overviews. Cockburn *et al.* (2008); Hornbæk & Frøkjær (2001) use thumbnails of a data set in conjunction with a detailed view as an overview and in contrast with other types of information seeking strategies. While the thumbnail of the dataset will show the scope and context of the detail being viewed, it is weaker on the other aspects. In acting as a guide towards areas of interest, there are cases where thumbnails do do this however, there are many cases where the thumbnail is too small to show anything of interest. In the case of the example thumbnail+detail (Figure 3.1), the thumbnail provides access to information pertinent to certain tasks – such as *where is Cote d'Ivoire?* – but is less likely to support more complex interaction. It is a simplistic view that does not encourage further exploration, nor does it give much of a sense of structure. For example, if the overview showed the outline of other countries, it would provide a sense of how Cote d'Ivoire compares in terms of size. Seeing the interrelation between the countries might lead a user to look at them in detail too.

One can imagine an interface where mousing over the thumbnail in Figure 3.1 would show the corre-

sponding detail in the larger image; the interface would provide easy navigation to different geographical locations but would be less effective in giving users a sense of what they might glean from the detail. The interaction is more driven by looking at the detail and understanding its context rather than looking at an overview and exploring a detail it highlights. This distinction sets the thumbnail apart from an overview.

Thumbnails used in cases such as in a slide sorter in a presentation application, provide an excellent mechanism for re-arranging the order of the slides but are less useful for other tasks such as describing the content of the presentation. Also, it is the distinctiveness of the slides that facilitate the effectiveness of the overview; when the slides are quite uniform, thumbnails provide little to identify the individual slides. In essence, they are a useful tool that could be used in an overview but on their own, they lack the sophistication of the semiotics that are built in to many overviews.

3.2 Changing modality: addressing auditory overviews

So far in this chapter, I have reviewed the research literature to identify what is said about overviews (Section 3.1.1) and then analysed a variety of visual overviews to assess what makes an overview (Sections 3.1.2 and 3.1.3). This section turns the focus back towards auditory overviews. I posit that auditory overviews only differ from visual overviews in modality and that the six characteristics will apply to them as well. To test this, I go over auditory overview literature with the findings from my analysis, as seen in Table 3.5.

3.2.1 Auditory overviews

Recent auditory overview researchers include Zhao, Kildal, Brown, Finlayson and Stevens. Zhao (2006; 2004; 2006; 2008) extend the Information Seeking Mantra (Shneiderman, 1996) to auditory interfaces in order to represent U.S census data, eschewing visual terminology by using the term *gist* instead of overview; Kildal's work (2006a) centres on overviews of numerical tabular data, such as in a spreadsheet; Brown (2006) provides overviews of edge/node graphs; Finlayson (2005) focuses on an overview of Java source code; while Stevens (1996; 1994; 1996) makes mathematical equations accessible through an *auditory glance*. Other research that touches on an auditory version of the Information Seeking Mantra is Pérez-Quiñones *et al.* who replace *overview* – again to avoid using visual terms for an auditory interface – with *situate*. The purpose was to provide voice interaction to web pages. While other auditory

work contain elements of overviews, these works specifically identify themselves as overviews. I assess each to determine how well the discovered characteristics of overviews applies to auditory overviews. The assessment is somewhat hampered by not having direct access to the overviews and thus relying on the authors' descriptions of their overviews; however, they are each described to have a majority of the characteristics, leading to the conclusion that the six characteristics hold for auditory and visual overviews alike.

U.S. census data

Zhao (Zhao, 2006; Zhao *et al.*, 2004, 2006) applies the information-seeking mantra to her work on auditory exploration of U.S. census data. The mantra becomes Auditory Information-Seeking Actions (ASIA). Here, in an attempt to distance herself from using visual language, the term *gist* replaces the term overview. A gist is a short audio clip that describes the detailed data. In her thesis, Zhao (2006) sets forth guidelines for the duration of a gist, its interaction and its latency. She argues that a gist should be no more than 10 seconds long due to the capacity of human short term memory, should be low latency (less than 100 milliseconds) and should be synchronised with other modalities to support multi-modal interaction. Only the proposed length of the gist is overview-specific; the other two guidelines pertain to general auditory interaction.

The overview of U.S. census data shows four of the characteristics: all of the major ones and one of the minor ones. The overview describes the scope by sonifying data points for all 50 states. The overview is also brief (less than 10 seconds) and acts as a guide by highlighting the variations in population across the United States. Context is heard by listening to neighbouring states. A listener can hear a snapshot of various census by selecting a different census year. It is less clear in the literature whether the overview exposes structure or encourages exploration though it is presumed that large changes in adjacent values might encourage exploration, as might perceived patterns. Table 3.12 shows how the overview of census data meets the attributes identified in Section 3.1.3.

	Scope	Guides quickly	Context	Structure	Encourages exploration	Snapshot
Census data	x	x	x	?	?	x

Table 3.12: The attributes and characteristics of census data.

Tabular data

Tabular data is data that is displayed on a grid. This sort of information is difficult to display in audio. In his work sonifying tabular data, Kildal & Brewster (2006a,b) do not formally address what an overview is. However, the authors state that in an overview, detail is irrelevant and that an overview can bring out patterns/trends in the data. One key aspect of the tabular data overviews is that they are row/column based. In other words, the overviews cover a subset of the data set. A user gains an understanding of the whole by comparing the row or column subsets. Using a stylus, the user iterates through the columns or rows, controlling the speed of the overview. The stylus interactions also allow the user to focus only on what they feel is relevant as opposed to the entire dataset. As the stylus travels over a column or row, it plays a representation of the numerical data contained within. Thus, Kildal & Brewster do not consider that overviews need to be comprehensive, merely that they represent a large enough subsection that a user can begin to locate salient features. In this case, the data is numerical and the overviews facilitate locating outliers: where numbers are especially high or low.

The tabular data overview shows evidence of all but one – a minor one – of the overview charac-teristics. The representation of all the columns/rows shows the scope of the detailed data. The brevity of the overview is determined by the user and is also driven by the number of rows/columns. However, the Sonification of each row/column is extremely brief and the overview as a whole can be considered short. The overview guides the users to columns/rows showing high or low values and shows context through contrast to neighbouring rows/columns. The structure is clear as the tabular nature of the data is intrinsic to the overview. The exposition of salient features, in this case high and low values, can encourage exploration. The playing of the overview is user-directed and as a result, it is less clear how well it could represent changes in states (i.e. historical snapshot). Table 3.13 shows how the tabular data overview meets the attributes identified in Section 3.1.3.

	Scope	Guides quickly	Context	Structure	Encourages exploration	Snapshot
Tabular data	x	x	x	x	x	?

Table 3.13: The attributes and characteristics of tabular data.

Edge/node graphs

Edge/node graphs are a way of representing interconnected data. For such graphs, Brown *et al.* (2006) aim to create an auditory equivalent to a glance. They set out two requirements for the overview:

to give an impression of size and complexity, and to describe the topology. The audio glance is an organised iteration through the graph that spreads from the left-most node: each node plays, then each node connected to it and so on. The basis for this is highlighting the relationships between the nodes and not the spatial layout.

The overview of edge/node graphs shows all but one – again a minor one – of the overview characteristics. The overview's main purpose is to describe entire graphs and their layout and thus it exposes the scope and structure of the detailed data. It shows context by describing the interconnections. While the size of the graph drives the length of the overview, it uses short non-speech sounds and is likely to be brief. The exposition of the features of the graph guide the listener through its layout and can provide a snapshot of the graph in various states. There is no evidence to the contrary, but it is hard to determine if the overview encourages exploration. Table 3.14 shows how the graph overview meets the attributes identified in Section 3.1.3.

	Scope	Guides quickly	Context	Structure	Encourages exploration	Snapshot
Edge/node graphs	x	x	x	x	?	x

Table 3.14: The attributes and characteristics of edge/node graphs.

Source code

Source code is computer programming code. It is plain text and very syntactically strict. Often, a single code file will have several thousand lines of code. Finlayson & Mellish's (2005) overviews describe Java source code. Similar to Brown *et al.* and Kildal & Brewster, they do not address what makes an overview. However, as with Kildal, the focus is on abstraction of the data and detailed data is obscured. The authors concentrate on the types of statements in the code rather than the statements themselves. The code is divided into three categories of statements, with nested statements exposed through a more complex representation. By describing nested statements, a user can perceive the hierarchical structure of the code. In other words, the overview is a broad iteration, describing the entire program and its structure.

This overview is based on Vickers & Alty (2005) who created auralisations of computer programs. They used musical constructs to aid novice programmers to identify bugs in Pascal code. The difference between their work and Finlayson's work is that the goal of Vickers & Alty's project, called CAITLIN, was to identify where problems occurred such as improperly terminated IF statements rather than pro-

viding a representation of the overall program structure.

Most of the major and minor overview characteristics are present in the overview of program source code. The overview shows scope and structure by summarising the code in order. Context is also apparent through this iteration through the code. Additionally, it is possible to have historical snapshots of the code as it evolves. What is less clear is if the overview is brief. By identifying various programming structures, the overview can act as a guide. However, the length of the code will drive the overview length and code source files can be several thousands of lines long. It is difficult to tell if the overview encourages exploration. Table 3.15 shows how the overview of computer code meets the attributes identified in Section 3.1.3.

	Scope	**Guides quickly**	**Context**	**Structure**	**Encourages exploration**	**Snapshot**
Source code	x	?	x	x	?	x

Table 3.15: The attributes and characteristics of program source code.

Mathematical equations

Mathematical equations are the language of maths. They provide complete descriptions of potentially complex relationships between variables and allow for their manipulation and analysis. When read aloud, they can easily and quickly become incomprehensible and misunderstood. Stevens (1996); Stevens *et al.* (1994, 1996) tackle the problem of making mathematics more accessible to visually impaired people. They created an auditory glance with the goal of expressing high-level structure to facilitate planning how to approach the mathematical expression. The auditory glance, which is an overview of the equation, describes the general shape of the expression and provides enough specifics to understand the complexity but the specific terms are obscured. For example, a user might hear that something was a number but not what that number is. The auditory glance would allow the user to understand perhaps that the expression is a quadratic equation and the user would need to explore further in order to hear the exact terms of the equation. They used algebraic earcons, composed with timbres, rhythm and prosody to describe the equations. The earcons describe the syntax of the expressionshowing aspects such as super/subscripts and describing the location of the various parts of the equation and their relative sizes. In essence, the auditory glance provides a framework for further exploring the equation. Experiments confirmed that participants could discern the complexity and shape of the equations and they were able to decipher the expressions while listening.

⇐item 15: tackles -> tackle

The type of information exposed in the glance shows what Stevens *et al.* (1994) prioritised for use in an overview. The specifics were not exposed but the intention was that participants could identify major segments of the expression through prosody and that the glance describes the entire expression. In terms of the characteristics identified through the visual overviews, the auditory glance satisfies all of them. The glance is comprehensive and shows the scope of the expression and describes the location of items through prosody. New items are distinguished through timing and pitch, allowing for context to be heard. The structure and syntax was proven to be discernable through user studies. While not the stated goal of the glance, it could be used to present a historical snapshot. For example, as a user worked with an expression and manipulated its shape, the glance could expose the changing shape by comparing two glances. The stated goal of the glance is to allow listeners to plan how they will approach the mathematical expression and thus it encourages exploration. Table 3.16 shows how this overview satisfies the overview characteristics.

	Scope	Guides quickly	Context	Structure	Encourages exploration	Snapshot
Mathematical equations	x	x	x	x	x	x

Table 3.16: The attributes and characteristics of mathematical equations.

Voice access to web pages

Webpages rely on their spatial layout and visual characteristics to guide the users. Various graphical elements draw the user's eyes to salient features. However, without a visual component, it is difficult to fluidly navigate the webpage. Here, an overview can facilitate navigation. Pérez-Quiñones *et al.* (2003) adapt the Information Seeking Mantra to guide the development of a voice system for accessing web pages. Because of the difference in auditory versus visual interaction, the authors propose *situate* instead of *overview* and describe it as a method to provide "an understanding of [the page's] structure" [p. 857] and to help users locate themselves within the information space. In other words, *situate* answers where the user is on the page and what options are available. The authors imply that this support enables quick navigation from one major section of the page to another. This is not an overview in the strictest sense; it is a system that exposes high level information. However, Pérez-Quiñones *et al.* focus on navigating the detail while occasionally accessing structural information. This is the opposite of some of the other work described here (e.g. Brown *et al.*; Finlayson & Mellish; Kildal & Brewster; Plaisant & Shneiderman; Stevens *et al.*, etc.) where the interaction starts with the overview which guides users to areas of interest.

The *situate* command in the voice access to web pages (Pérez-Quiñones *et al.*, 2003) is similar to an overview. Its stated goal was to facilitate navigation and expose the structure of webpages. As such, it appears to have several of the overview characteristics. As a navigation tool, it encourages exploration and acts as a brief guide to the page. It is unclear if the overview is semantic or simply structural. This means it is difficult to tell if context is exposed or snapshots of the page are possible. Scope is also difficult to determine. Table 3.17 shows how the *situate* command meets the attributes identified in Section 3.1.3.

	Scope	**Guides quickly**	**Context**	**Structure**	**Encourages exploration**	**Snapshot**
situate	?	?	x	x	x	x

Table 3.17: The attributes and characteristics of the *situate* command.

3.2.2 How the discovered characteristics fit auditory overviews

Table 3.18 shows a summary of all the auditory overviews from the literature and how they satisfy the attributes identified through my survey of overviews. With the visual overviews (Section 3.1.2), there was a clear demarcation between major and minor characteristics. Showing scope, context and quickly guiding all emerged as important characteristics. The remainder (showing structure, encouraging exploration and providing a historical snapshot) were not omnipresent and thus are minor characteristics. This pattern is not repeated with the auditory overviews. Only showing context appeared definitively in all the overviews. Showing scope, structure and providing a historical snapshot were quite frequent while quickly guiding and encouraging exploration were harder to show in the overviews.

The use of the word *definitively* is key here. The overviews surveyed in Section 3.1.2 were all easily generalised due to their prevalent natures. Examples of tables of contents, file managers, line graphs and abstracts are plentiful. This is not the case with the auditory overviews where I am reliant on the quality and the comprehensiveness of the written descriptions of original research. This makes it difficult to properly assess them in the same way as visual overviews. As such, this review of auditory overviews is conservative as to which attributes match and which do not.

With this caveat in mind, Table 3.18 shows that the auditory overviews each match four or more of the six overview attributes. This indicates that the overview attributes do apply to auditory overviews. As far as which attributes apply to auditory overviews, encouraging exploration is the weakest, with only two specifically mentioning navigation and exploration. Encouraging exploration is also the hardest

	Scope	Guides quickly	Context	Structure	Encourages exploration	Snapshot
Census data	x	x	x	?	?	x
Tabular data	x	x	x	x	x	?
Source code	x	?	x	x	?	x
Edge/node graphs	x	x	x	x	?	x
situate	?	?	x	x	x	x
Mathematical equations	x	x	x	x	x	x
	83%	66%	100%	83%	50%	83%

Table 3.18: **The attributes and characteristics of overviews from the literature.** The final line of the table shows how many of the overviews reviewed showed evidence of each attribute.

attribute to prove since overview researchers do not address the issue. Guiding quickly is the weakest of the major characteristics, turning up in only four of the six auditory overviews. Overall, the strongest statement that may be made is that the overview attributes seem to apply to auditory overviews as well however, this cannot be proved conclusively.

3.3 Conclusion

While not all the components described in Table 3.5 will be present in all overviews, these cover the major characteristics of an overview. As with an abstract, where the structure and organisation of the article is not addressed, an overview should draw upon those characteristics that will best serve its purpose. In other words, the definition of an overview must be fluid in order to adapt to the material that is summarised. Purpose also plays a role as shown by the differences between an abstract and a table of contents (3.1.3) and the discussion of task dependency in Section 3.1.3.

The main purpose of an overview is to quickly guide to the detailed data. This means it should be a brief but complete and sufficiently accurate representation of the data set. The overview also manages the user's expectation by delimiting the data set – where does it start and end – and exposing patterns and relationships – how it is organised – between the various subsets of the data. The power of an overview lies in exposing information that is not easy to grasp by perusing the detailed data.

An overview is a short, abstract view of the detailed data that rapidly exposes high level information. This information in turn exposes salient features and interrelationships, thereby engaging the user's curiosity. Its usefulness comes from providing a shortcut to in depth knowledge that could otherwise only come from extreme familiarity.

The above general guidelines are a product of my analysis of existing research and of a variety of

overviews. In Section 3.1.1, I looked at the use of overviews in information-seeking and other sources. These references do not formally define "overview", but do generally agree that overviews provide guidance, facilitate navigation and are complete representations of the detailed information. I used these as the basis for my analysis of the following overviews: tables of contents, file managers, line graphs and abstracts (see Section 3.1.2) which, in turn, confirms the initial assumptions and gives a deeper understanding of the role overviews play. I then assessed the auditory overviews (Section 3.2.1) to see how the attributes apply to auditory overviews. Given the variability in the way the research studies of auditory overviews are reported, it is difficult to show that all the overview attributes and characteristics identified in my analysis of commonly used overviews apply to all the auditory overviews discussed in this chapter. However, as the discussion of each auditory overview shows, each one does individually indicate the presence of most of the overview attributes. In conclusion, this in depth analysis of overviews provides a set of guidelines for what an overview, particularly visual and auditory, should aim to do.

Chapter 4

Augmenting an auditory application with an overview

In the previous chapter, I analysed several types of visual overviews and reviewed auditory overviews in the literature. The result of the analysis helped define the characteristics of an overview. However, it is important to remember that overviews are part of a bigger whole; for example, a table of contents is part of, and supports the exploration of, a body of text such as a book. This chapter and the next describe an auditory application and an experiment to understand how the inclusion of an overview changes the user's perception of the detailed information and how this can further clarify the role of overviews in an auditory interface. This chapter covers the auditory application and the design of the experiment while the Chapter 5 reports on the results. The two chapters are intended to be read together.

The auditory application that I developped to examine interaction with or without an overview is an auditory public transportation system application. Public transportation systems are practical, real-world examples of non-trivial, interconnected graphs. In a speech system, the interface can describe one transportation line at a time and it is up to the user to deduce the relationship between the lines. It is a challenge to auditorily represent such information in a clear and concise manner. The idea behind this choice of application was that complex detailed information and associated tasks would highlight differences in usage more fundamentally than a simpler one would by making users more reliant on the information provided in the overview.

I used the complexity of the application to draw out how participants performed route-finding tasks with and without the overview to guide them. In the experiment, participants used the Speech-Only Interface (SOI) as well as the Overview-Augmented Interface (OAI). The overviews in the application

are short, non-speech sounds that described the entire metro system as well as subsets there of. The experiment and the application were specifically designed to draw out how much participants were able to gain an understanding of the metro system and to observe how participants used the underlying data differently when supported with an overview. Chapter 5 presents the results of this experiment.

4.1 Pairing overviews with detailed information

4.1.1 Stand-alone overviews

On their own, overviews can provide clues to the content of what they describe, the organisation of the content and some key details. In a table of contents, the linear flow of the written material is visible. With line graphs, the spatial layout and connections between data points are visible. In isolation, each of these overviews provide summary information in a manner that is useful. And they can be used standalone in certain scenarios. For example, the table of contents might be sufficient in deciding whether to purchase a book or article. This can be seen in the real world: Amazon provides the "search inside" feature[1] so that users of their site can see the table of contents as well as an excerpt from the book. In the academic world it is frequent for online libraries, such as the ACM Digital Library[2] or Elsevier[3], to publish abstracts for free, while access to the full text is for purchase or comes as part of a membership. In these models, the overview is considered sufficient for deciding interest.

4.1.2 The purpose of studying overviews paired with detailed information

Overviews are most beneficial when used in conjunction with the information they describe. A table of contents shows where to find a section or paragraph on a particular subject matter; or it can pinpoint a moment in a narrative that the reader wishes to revisit. It serves as a guide to the actual content. Likewise, a map serves as a guide to a location but the details of the journeys within that location itself cannot be contained in the map. Chapter 3 highlighted that using an overview with the detail can result in better organised detailed data (see Section 3.1.3). The next question is whether the existence of an overview has an effect on how people use the detailed information. While there has been work on what information people should get from overviews (see Section 3.1.1), there is little on how the use of an overview influences user interaction with the detailed information.

---

[1]http://www.amazon.com/gp/help/customer/display.html?nodeId=10197041
[2]http://portal.acm.org/dl.cfm
[3]http://www.elsevier.com/

There are many questions that arise about the interaction between detailed information and an overview. Or rather, there are questions about what role overviews play in the user experience. How does the inclusion of an overview influence user behaviour? Can inclusion of an overview change the form of the underlying information? For example, can viewing chapter and section headings in a table of contents sway the author into changing them? Or the flow of the document? Can the creation of an overview create a feedback loop where it impacts the structure? Is the work perceived differently because of the presence of an overview? The relationship between the overview and the detailed information has not received much study.

The study described in this chapter is a beginning in this line of research.. As a first step, I focused on the differences in user behaviour, depending on the presence/absence of an overview, in order to understand the impact of an overview on the user interaction in the context of an auditory application.

4.2 Exploring the effect of overviews: an auditory public transportation system application

The goal of this experiment was to explore overviews in their context of use. That is, I studied how an overview is used in the presence of detailed information. I developed an auditory application that allows users to explore a public transportation system. When used with the SOI, users navigate to find detailed information, which is read out to them. When the application is augmented with overviews, users hear a non-speech representation of the public transportation system which they can manipulate to customise what they hear. I chose to represent a public transportation system because it would be a familiar construct, yet has complex interconnections that are difficult to express in speech and non-speech alike. Participants in the experiment were assigned route finding tasks that would make it necessary for them to understand the complexity of the transportation system. This setup, where I can turn off overviews, allowed me to observe how participants interacted differently with the application depending on whether they had access to overviews or not.

4.2.1 The auditory public transportation system application

The auditory public transportation system application is primarily speech-based with the option to add on overviews. I will first start by describing how the application works in its speech-only form and then how it is augmented with overviews. For details on its design rationale, see Nickerson *et al.* (2007) and

Nickerson (2007).

4.2.2   The speech-only application

The application is divided into two main parts. In one part, users explore the public transportation system via the lines that run through the system. This is called the Legend. In the other part of the system, users can explore via the stations. This is called the Index. This provides users two views of the system and, as they explore, the various pieces of information are read out to them.

Navigation takes the form of a tree which participants traverse using the cursor keys. From the root, users can access either the Legend or the Index. In the Legend, the application lists the names of each line in the transportation system. From each of the lines, users can explore the stations on that line in the order that they appear in the line, i.e. in travel order. And in turn, users can explore what (other) lines go through that station. On the Index side of the application, the same organisation is repeated: the Index lists all the stations in the transportation system in alphabetical order. From each station, users can explore what lines go through them and in turn, from each line, what stations are on those lines. Under each station or line node in the tree, there is an Options level where users either choose to explore the current line or station, or to navigate to the parent node (the line or station node) or the root node. The diagram in Figure 4.1 shows the organisation of the application tree, using an example transportation system shown in Figure 4.2. This is a simplified system for the purposes of explanation. Figure 4.8 shows the system that the participants actually used in the experiment.

The interaction with the application is via the computer keyboard. Participants use the cursor, <home> and <enter> keys to navigate through the application tree. To select a node, and thereby navigate to its first child or to apply a particular option, participants can use either the <down> arrow or the <enter> key. In order to go to the root of the application, participants can use the <home> key. Table 4.1 shows an example interaction with the application, using the example transportation system shown in Figure 4.2.

Using the application, participants can look up stations via the Index and find out what lines they are on. They can also explore lines via the Legend, list the stations on each line and find out if a station is a transfer station by exploring what lines go through it. This gives the participants sufficient tools to plot out journeys through the public transportation system. For example, given two stations, participants can look up the stations in the Index, find out what lines they are on and then explore each of those lines to find a transfer station that connects the lines in question.

APPLICATIONROOT | SELECTVIEW | INITIALSTATIONAND LINELISTS | OPTIONS | SECONDARYLINEAND STATIONLISTS | OPTIONS | TERTIARYSTATIONAND LINELISTS

Root

Index

Foxtrotstation — Gohome, Goback, Linesthrough Foxtrotstation → TheBlueline → Gohome, Goback, StationsontheBlueline → Foxtrotstation, Alphastation, Echostation, Deltastation

Echostation — Gohome, Goback, Linesthrough Echostation → TheBlueline → Gohome, Goback, StationsontheBlueline → Foxtrotstation, Alphastation, Echostation, Deltastation

Deltastation — Gohome, Goback, Linesthrough Deltastation → TheBlueline → Gohome, Goback, StationsontheBlueline → Foxtrotstation, Echostation, Alphastation, Deltastation

Charliestation — Gohome, Goback, Linesthrough Charliestation → TheRedline → Gohome, Goback, StationsontheRedline → Charliestation, Alphastation, Bravostation

Bravostation — Gohome, Goback, Linesthrough Bravostation → TheRedline → Gohome, Goback, StationsontheRedline → Charliestation, Alphastation, Bravostation

Alphastation — Gohome, Goback, Linesthrough Alphastation → TheRedline → Gohome, Goback, StationsontheRedline → Charliestation, Alphastation, Bravostation; TheBlueline → Gohome, Goback, StationsontheBlueline → Foxtrotstation, Echostation, Alphastation, Deltastation, TheRedline

Legend

TheRedline — Gohome, Goback, StationsontheRedline → Charliestation → LinesthroughCharliestation, Goback, Gohome; Alphastation → LinesthroughAlphastation → TheRedline, TheBlueline; Goback, Gohome; Bravostation → LinesthroughBravostation → TheRedline; Goback, Gohome

TheBlueline — Gohome, Goback, StationsontheBlueline → Foxtrotstation → LinesthroughFoxtrotstation → TheBlueline; Goback, Gohome; Alphastation → LinesthroughAlphastation → TheRedline, TheBlueline; Goback, Gohome; Echostation → LinesthroughEchostation → TheBlueline; Goback, Gohome; Deltastation → LinesthroughDeltastation → TheBlueline; Goback, Gohome

Figure 4.1: **The structural organisation of the public transportation system application.** Please note that this diagram has been rotated 90° for clarity; in the application, participants moved up and down to change levels, not left and right.

LEVEL 1: the root of the application where participants find themselves when they start the application.

LEVEL 2: choose to explore the system by lines (Legend) or stations (Index).

LEVEL 3: list the lines (Legend) and the stations (Index) in alphabetical order.

LEVEL 4: (options level) choose to explore the stations on a line (Legend) or explore the lines through a station (Index); also two navigation options: go back (parent node) or go home (application root).

LEVEL 5: stations on a line in travel order (Legend) or lines through a station in alphabetical order (Index).

LEVEL 6: a level of options similar to level 4.

LEVEL 7: lines through a station in alphabetical order (Legend) or stations on a line in travel order (Index).
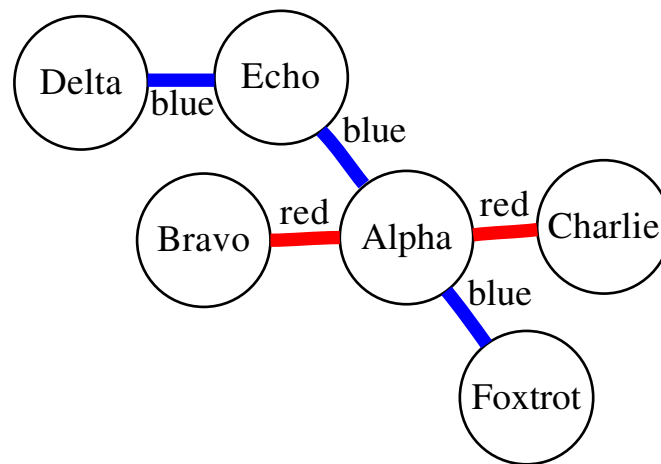
Figure 4.2: **Example public transportation system map for the purposes of explaining the application.** This is a simple six-station transportation system as used in Figure 4.1 and in Table 4.1. There are two lines: the Red and Blue lines. Alpha station is the transfer point for the two lines. The Red line stations in order are Bravo → Alpha → Charlie; the Blue line stations are Delta → Alpha → Echo → Foxtrot.

4.2.3 The application augmented by overviews

The OAI uses the SOI as a base. As such, the speech and commands are the same. The main difference is that the OAI has non-speech sounds that help describe the transportation system. There are three types of overviews in the interface: a Main overview, Line overviews and Station overviews. The line overviews, which are not customisable, describe the scope of the line in the transportation system – i.e. where it starts and ends relative to the entire system. The station overviews, also not customisable, describe where the station is in the transportation system. This section describes the overviews themselves.

The main overview

The main overview describes the entire transportation system. This overview loops and plays continually throughout usage of the application. It does, however, go silent whenever any speech occurs. Participants can customise this overview by setting filters on the lines and stations they are interested in. The unfiltered overview is quite short and, in the case of the public transportation system used in the experiment, lasted 5.9seconds.

The overview is made up of several different sounds. First, each line in the transportation system has its own frequency. The frequencies are logarithmically spread out to ensure a maximum perceptual difference of pitch between them. Second, there is a background sound made up of white noise that oscillates in amplitude during each loop of the overview such that it is loudest at the beginning and the end, and quietest in the middle of the overview. This is in order to help spatially localise the various

Location	Key pressed	Sounds heard	Result
Application root	<down> arrow	*changing level* sound, "the Legend"	focus is now on the first child of the application root: *the Legend* node
Legend	<right> arrow	"the Index"	focus is now on the next sibling of the *Legend* node: *the Index* node
Index	<right> arrow	*looping* sound, "the Legend"	there is no sibling node to the right, focus loops around back to *the Legend* node
Legend	<enter> key	*changing level* sound, "the Blue line"	focus is now on the first child of *the Legend* node: *the Blue line* node
The Blue line	<down> arrow	*changing level* sound, "stations on the Blue line"	focus is now one the first child of *the Blue line* node: the *stations on the Blue line* node
Stations on the Blue line	<right> arrow	"go back"	focus is now on the next sibling of the stations on *the Blue line* node: the *go back* node
Go back	<enter> key	*changing level* sound, "the Blue line"	focus is now on the parent of the *go back* node: *the Blue line* node
The Blue line	<right> arrow	"the Red line"	focus is now on the next sibling of *the Blue line* node: *the Red line* node
The Red line	<down> arrow	*changing level* sound, "stations on the Red line"	focus is now on the first child of *the Red line* node: the *stations on the Red line* node
Stations on the Red line	<enter> key	*changing level* sound, "Bravo station"	focus is now on the first child of the *stations on the Red line* node: the *Bravo station* node, the first station on the Red line

Table 4.1: **Example interaction with the public transportation system application.** A participant navigates from the root of the application down to Bravo station which is on the Red line. The *sounds heard* column shows what the participants would hear, both non-speech and speech. The *key pressed* column shows what keyboard keys were used to perform an action.

sounds in the overview. Each of the lines in the public transportation system is represented by a series of beeps at the frequency of that line. So in essence, when the overview starts, participants hear the shh-ing white noise sound that indicates the beginning of the overview and then a series of beeps. Gradually the shh-ing sound decreases and then increases as it reaches the end of the overview.

The series of beeps that make up each line represent each of the stations on that line. It is as though the overview were scanning the lines west to east. The end effect is that the lines are interleaved and the timings of the stations are based on the relative location to one another. Transfer stations are represented by a composite sound of all the lines that go through them. In the case of the simplified transportation system described in Figure 4.2, you would hear the stations in the following order: Delta → Bravo → Echo → Alpha → Foxtrot → Charlie, where Bravo and Charlie have the Red line sound, Delta, Echo and Foxtrot have the Blue line sound and Alpha is a composite Red/Blue sound. The shh-ing white noise
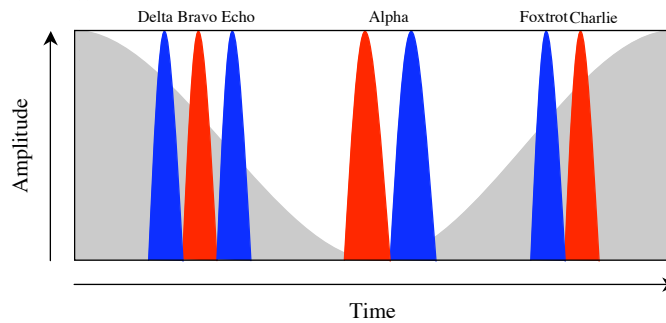
Figure 4.3: **The main overview.** This shows the timings of an unfiltered main overview, using the simple six-station system described in Figure 4.1. The y-axis shows amplitude and the grey curve in the background shows the levels of the background shh-ing noise which starts loud, is quiet in the middle and loud again at the end of each overview loop. The x-axis shows time and the red and blue represent the respective sounds for the Blue and Red lines. Delta is presented first, then Bravo, Echo, Alpha, Foxtrot and Charlie. The timings for Delta, Bravo, Echo, Foxtrot and Charlie are based on the fact that they are simple stations – only one line goes through them – while Alpha as a transfer station has the sounds of each of its lines, padded by silence on each side to separate it from the presentation of the simple stations.

sound is loudest just before Delta and after Charlie; it is quietest just as Alpha plays. Simple stations – ones that have only one line going through them – are represented quickly (0.045s) by the sound corresponding to its line. Transfer stations are represented by the different sounds of its lines played a bit more slowly and padded by silence on each side to differentiate it from simple stations (0.12s padding on each side, 0.06s for each line that goes through it). Figure 4.3 shows a graphical representation of what this sounds like.

As mentioned above, this main overview can be filtered. The controls for filtering are located on the options level in the application tree. Figure 4.4 shows what the modified actions look like. The first option remains the same (*stations on [...] line* or *lines through [...] station*). The filtering options appear directly after that. The first filtering option allows participants to put a filter on that line or station (*filter on [...] station* or *filter on [...] line*). The second filtering option allows participants to cancel any filters (*clear all filters*). The last two options are the same as in the SOI.

Line and station filters work by highlighting the lines and stations in the main overview. These filters can be combined. When participants set their first line filter all other lines fall silent. All that can be heard is the one line with the filter set. There are several pieces of information that participants can hear with a single line filtered. They will be able to hear the extent of the line in the system from when they hear the first sound – the first station – of that line to when they hear the last sound of that line – the last station. They can also hear how many transfer stations are on that line and where those transfer
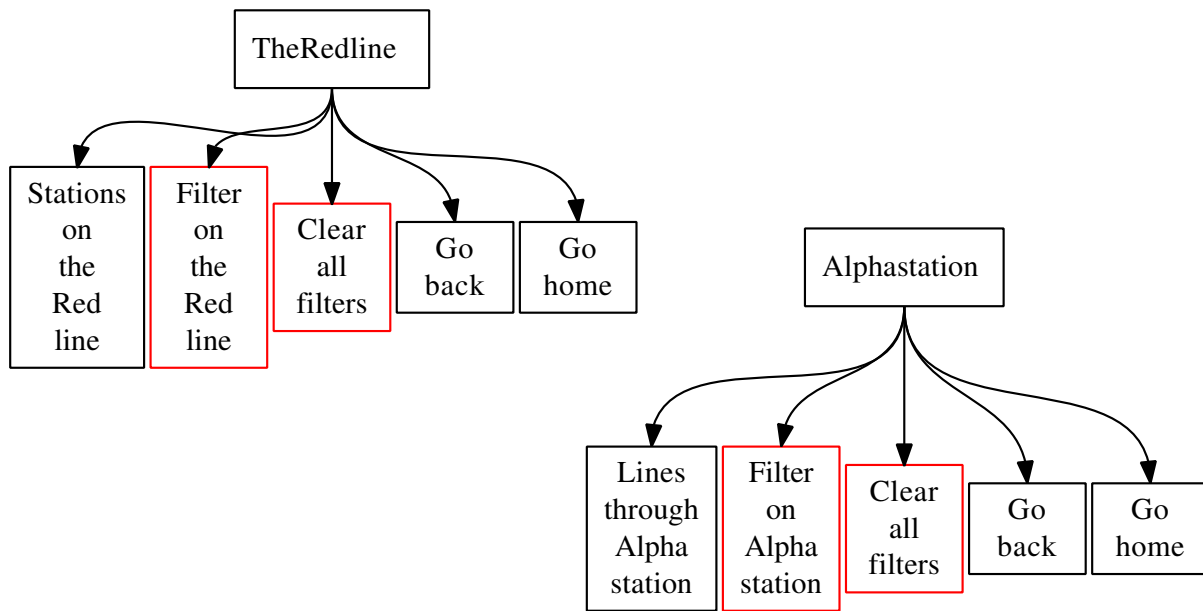
Figure 4.4: **The additional filter options in the OAI.** Between the *Lines through/Stations on [...]* node and the *Go back* nodes, there are two filtering options: the first *filter on [...]* allows participants to set a filter; it toggles to *clear the filter on [...]* so the individual filters can be turned on and off. The second option, *clear all filters*, allows participants to remove all filters, or rather to reset the main overview.

stations are. Each transfer station contains not only the sound of the line filtered but also the sounds of the other lines that go through that station. Those other sounds stand out from the filtered overview. The appropriate use of filters allows users to hear what are the intersecting lines. When they set their second line filter, they hear both lines and any transfer station between those two lines is highlighted.

When participants add a station filter, that particular station is highlighted in the main overview. The overview slows as it approaches the station of interest, a spearcon of the station's name plays and then the overview speeds up again. This is where the background shh-ing white noise becomes particularly important: as the overview is time-stretched in places, the audibility of that background sound indicates how far along the overview is. What a station filter tells participants is where sequentially that station is in the transportation system, what lines go through it and the station's name. The station name is spoken very quickly and thus is most intelligible when the user is familiar with the station name through other interaction, e.g. because s/he explicitly set a filter on it or encountered it through exploration of the application. When multiple line filters trigger a station filter, participants can hear where and how often the lines intersect.

In other words, the main overview can be manipulated to display interconnections. By setting filters on the lines that run through two stations, participants can hear a potential transfer station that will allow

Goal	Step	Key pressed	Location	Navigation sounds	Line overviews	Station overviews	Main overview
Find out what line Bravo is on	1	<down>	Index	*change level*			The Red and Blue lines — no station highlighted
	2	<right>	Legend				
	3	<down>	Alpha station	*change level*		Alpha station	
	4	<right>	Bravo station			Bravo station	
	5	<down>	Lines through Bravo station	*change level*			
	6	<down>	The Red line	*change level*	Red line		
	7	<right>	The Red line	*looping*	Red line		
Set a filter on the Red line	8	<down>	Stations on the Red line	*change level*			
	9	<right>	Set a filter on the Red line				
	10	<enter> or <down>	Clear the filter on the Red line (unsaid)				Only the Red line — no station highlights
Set a filter on Bravo	11	<up>	The Red line	*change level*	Red Line		
	12	<up>	Lines through Bravo station	*change level*			
	13	<right>	Set a filter on Bravo station				
	14	<enter> or <down>	Clear the filter on Bravo station (unsaid)				
Find out what line Delta is on	15	<up>	Bravo station	*change level*		Bravo station	The Red line — Bravo station highlighted
	16	<right>	Charlie station			Charlie station	
	17	<right>	Delta station			Delta station	
	18	<down>	Lines through Delta station	*change level*			
	19	<down>	The Blue line	*change level*	Blue line		
	20	<right>	The Blue line	*looping*	Blue line		
Set a filter on the Blue line	21	<down>	Stations on the Blue line	*change level*			
	22	<right>	Set a filter on the Blue line				
	23	<enter> or <down>	Clear the filter on the Blue line (unsaid)				The Red and Blue lines — Bravo and Alpha highlighted

Table 4.2: **Example usage of filters in public transportation system application.** An example of how a participant might use filters given the task of planning a route between Bravo and Delta stations. This example uses the network shown in Figure 4.2. The *goal* column shows what a user is trying to do and the *step* column details the numbers of the steps that the participant went through to manipulate the overview. The *key pressed* and the *location* columns show what action the participant took and where they ended up in the interface. The last four columns show what the participant hears.

Several things to note: filtering on the line first (steps 8 through 10 and steps 21 through 23) allows a participant to situate the station better than had the station been filtered first; in step 23, the participant applies a filter to the Blue line and the interface automatically highlights the transfer station between the Red and Blue lines – the participant can now hear that Alpha is the transfer station that allows travel between Bravo and Delta; in steps 6 and 7, and again in steps 19 and 20, the participant repeats the line information so as to confirm that there is only one line through each of the stations in question.
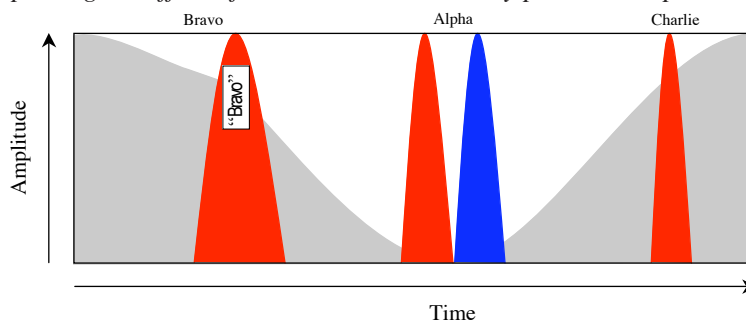
Figure 4.5: **The main overview filtered.** The state of the main overview here corresponds to step 15 in Figure 4.2. There is a filter on the Red line and Bravo station. The participant can hear all the stations on the Red line (Bravo, Alpha and Charlie) and Bravo is highlighted. The timings of Bravo and the silence before it are stretched to allow more emphasis. A spearcon of the station name "Bravo" is spoken while its sound is playing. Note that the Blue line sound still appears in the presentation of Alpha station. This allows participants to hear when there are transfer stations.

them to travel between those two stations. By adding subsequent filters on those two initial stations, participants can hear approximate distances between the various stations and in what direction they would have to travel to make the journey. Table 4.2 runs through a example scenario showing how the filters can be used and Figure 4.5 shows a filtered main overview.

Line filters do not affect the timing of the main overview, however station filters can stretch the overview by several seconds. When a station filter is set, that station and the four previous stations are affected. The overview is gradually slowed and the length of the filtered station and four preceding stations are increased by 25% to 125%.

Overview characteristics   The main overview is the most complex overview in the application. Table 4.3 shows how it satisfies the attributes that emerged from my analysis in Chapter 3. The main overview and the sub-overviews created by filters contain all of the attributes in its design. The shh-ing white noise in the background shows the limits of the metro system and thus the scope of the system. The stations are all represented, making the overview comprehensive. The display of transfer stations describes the structure and its interrelationships. It also facilitates understanding of the context of each station. As it is a static system, the attribute of providing a historical snapshot is less relevant, however, were there several evolutions of the metro, one could easily compare the overviews of each. While it might be difficult initially to hear the flow of a single line, usage of the filters, which isolate lines and highlight stations make the main overview more comprehensible. This is akin to visually encountering a new map, becoming familiar with it and then being able to identify features more readily at a glance. The filters and automatic highlighting are arguably creating sub-overviews. These sub-overviews quickly
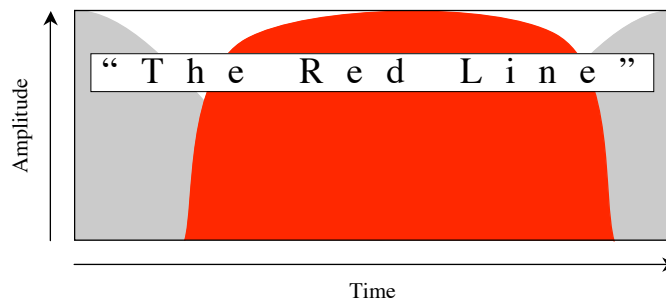
Figure 4.6: **A line overview.** The sound of the Red line starts when the first station of the Red line occurs and ends when the last one occurs. The entire overview is timed to take up the approximate duration of the speech "The Red line."

guide and encourage exploration, as shown by the changed interaction when participants used the OAI (see Section 5.1). As such, the main overview satisfies these qualities as well.

Line overviews

Line overviews play when participants navigate to a line node. The overview is based on the main overview in that it has the same background white noise but it is much shorter. A single tone in the sound of the line in question plays during the overview, defining where the first and last stations of that line are in the transportation system. It also has the added benefit that it reminds participants of what is the particular sound of that line. This is a static overview and cannot be manipulated. It is timed to coincide with the duration of the speech which accompanies it: the name of the line. In the experiment, the average time to speak each line name was 1.25s and the line overviews were scaled to fit into this time period. Figure 4.6 shows a line overview of the Red line using the example system from Figure 4.2.

Overview characteristics   The line overview is the simplest of the overviews in the system. Its main role is to identify the sound of the line and to describe where the line fits in the metro system. From the overview, it is possible to hear where the line starts and ends and how long a line it is. As such, the

	Scope	Guides quickly	Context	Structure	Encourages exploration	Snapshot
Main overview	x	x	x	x	x	x
Line overview	x	x	x		x	
Station overview	x	x	x	x	x	

Table 4.3: **The attributes and characteristics of the overviews used in the experiment.** Each of the overviews have the major overview characteristics (scope, guiding quickly and context). The line overview is the weakest one, with only one of the minor characteristics represented.
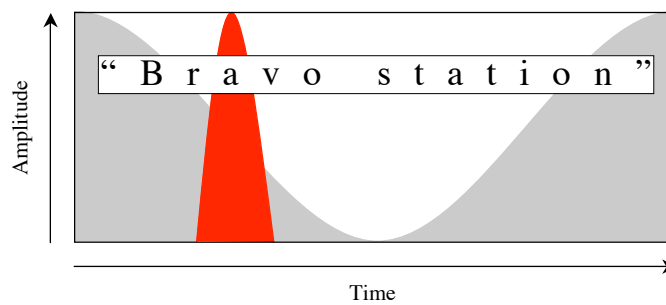
Figure 4.7: **A station overview.** This overview is timed to last approximately as long as the speech ("Bravo station"). The timing of the sound of Bravo station, here represented in red, comes from where it would occur in the main overview. Bravo station is at the beginning of the main overview and thus occurs early on in the station overview. Likewise, if this were an overview of Alpha station, it would play a Red line and a Blue line sound in the middle of the overview.

overview conveys the scope of the line and a certain amount of its context. To a lesser extent, it can quickly guide – e.g. where the user is looking for a particular line by sound, the overview identifies it as a line of interest. It can also provide a historical snapshot given several versions of the line or metro system (a change in system size would affect the overview). However, the overview is less effective in conveying the line's structure and does little to encourage exploration. Thus, the line overview displays all three of the major overview characteristics from Chapter 3 and one of the minor ones (see Table 4.3).

Station overviews

Like line overviews, station overviews are static and they play when participants navigate to a station node. Each station overview contains the same background white noise and a single beep indicates where the station is in the transportation system. The beep uses the sound of the line that goes through that station; if the station is on multiple lines, the sound that indicates the location of the station is a composite set of beeps using the sound of each line that goes through it. This overview allows participants to hear where a station is and match that up with what they hear in the main overview. The station overview also indicates whether the station is a transfer station or not without having to be explored. The timing of the overview (1.5s) is based on the speech that accompanies it: the name of the station. It is slightly longer than the line overview as it generally takes longer to say the station names than it does to say lines. Figure 4.7 shows the overview for Bravo station.

Overview characteristics   Station overviews provide information about where the station is in terms of the metro system and what lines go through it. The combination of the two provides the context and the scope of the station; the latter also shows the structure of the station in terms of the map: whether it is

a simple or transfer station. The sounds identify the lines that go through it, quickly guiding the user to information of interest. Were there several versions of the system, the overview would also provide a historical snapshot. As it is a simple overview – the detail it covers is quite small – it does less to encourage exploration. This overview displays all but one of the characteristics identified in Chapter 3 (see Table 4.3).

4.3 Design of the experiment

The participants used the public transportation system application in a route finding task. A practical task such as this pushes participants to gain more familiarity with the transportation system where a more open-ended exploratory one would not. I assigned participants a composite task that I call the **meeting up task**. In this task, I gave participants three starting stations. Their goal was to find out where these stations are in the transportation system and determine a mutually convenient meeting place for three fictional people coming from these stations. In order to perform this task, participants need to at least loosely map out the journeys between the three stations.

4.3.1 Hypotheses

Working from the basis that an overview acts as a guide, it seems logical that the presence of an overview will affect the usage of the detailed information. However, there is very little in the literature concerning what that effect might be. Most studies look at whether the overviews can impart information and rarely discuss the combined use of overviews with the detailed information to which they refer, let alone study the effect of the overview's presence on the interaction with the detailed information. The experiment described in this and the following chapter is primarily exploratory. The aim is to find out how the addition of an overview affects the user experience, with the implication that if we understand the impact of overviews better, they can be exploited more effectively in interfaces. The six characteristics outlined in Chapter 3 act as a framework for predicting how participants interaction will change in the presence of an overview.

1. **Scope of the metro system** Each of the overviews in the application is designed to give the user an idea of where lines and stations are in the larger picture. The expectation is that the participants will have a better idea of the size and complexity of the system.

2. **Quickly guides to information** The overviews should give a general sense of where things are, so that instead of searching for information, users can go directly to it. This indicates that participants will navigate more efficiently to pertinent information.

3. **Shows contextual information** The overviews present where stations and lines are located, while the filters highlight what is near. Participants will, as a result, have a better sense of, for example, how far apart stations are on a line and better judge distances.

4. **Exposes the structure of the metro system** The overviews strive to emphasise interconnections in the metro system. Participants will grasp how the metro lines interrelate.

5. **Encourages exploration** The overviews highlight features of the metro system that are of interest. Participants will explore more aspects of the system to satisfy their curiosity.

6. **Shows the state of the metro system** Since the metro system was static, the effect of this characteristic will not be seen.

These predictions fall into two major hypotheses. The first is that participants using the overviews will have a better understanding of the metro system in general. The second is that the overviews will change how participants use the menu system.

Understanding the metro system

To test participants' understanding of the metro system, they are interviewed after using each interface. Participants using the OAI are expected to know more about transfer stations within the system, have better knowledge of the lines and the size and layout of the system. This ties in with predictions 1 and 4 above. Tasks completed with the OAI should have better scores as they will have access to more contextual information (prediction 3).

Use of the menu system

Predictions 2 and 5 imply that tasks completed with the OAI will be faster since they will be guided to information of interest. Also, the exploration patterns should be broader since participants will be aware of more stations to investigate. This will be primarily apparent by looking at where and how participants navigate.

4.3.2 The tasks

The public transportation system

For the purposes of the experiment, I generated a fictitious public transportation network. Like the example network shown in Figure 4.2, the public transportation system used colour names for the lines and words from the phonetic alphabet for station names. The use of the phonetic alphabet was to ensure spoken names would not be easily confused. There are five lines in the transportation system: the Blue, Green, Purple, Red and Yellow lines. Each line has between 15 and 17 stations and crosses all the other lines. There are 73 stations in the transportation system, of which 8 are transfer stations. There is one transfer station with three lines through it (Lima station), the rest of the transfer stations being an intersection between only two lines. This transportation system can be seen in Figure 4.8. The system was designed to be complex enough to ensure that participants were using the auditory information as I explain further in Section 4.3.3.

Experimental procedure

As the main goals of this experiment were to study how the OAI affected understanding of the transportation system and information-seeking patterns in the application in comparison to the SOI, I chose a *within subject* design. Half the participants used the SOI first and the other half used the OAI first. Given that the application is novel and the auditory interaction is unfamiliar to the participants, a within-subject design allowed me to more easily cross-compare the usage of the two interfaces, eliminating errors introduced by having two distinct populations. Each participant performed four meeting up tasks, the first two with one interface and the second to with the other. Two hours were allocated for each experiment. Here is how a typical experiment ran:

Recruitment: I sent out a recruitment circular (see Section A.1) to students in the School of Electronic Engineering and Computer Science, the text of which was approved by the Queen Mary Ethics Board.

Welcome: upon arrival, participants read through the information sheet (see A.2); I explained the experiment to them; they were encouraged to ask questions and then we signed the consent form (see Section A.3).

Training: participants listened to a presentation (see A.5) detailing the context of the experiment, the interface they would be using and the tasks they would perform. Then, I ran through a simple
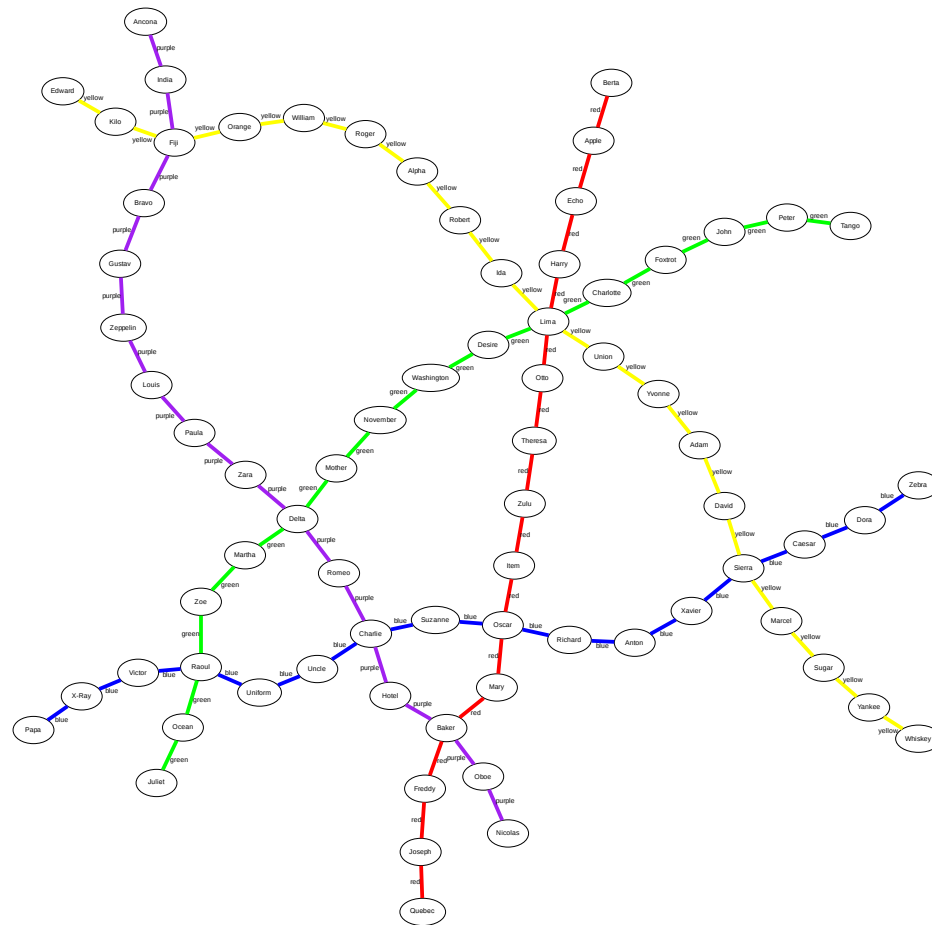
Figure 4.8: **Map of the public transportation system used in the experiment**. There are five lines in this transportation system: the Blue line, the Green line, the Purple line, the Red line and the Yellow line. Each line intersects every other line. There are eight transfer stations: Baker station, Charlie station, Delta station, Fiji station, Lima station, Oscar station, Raoul station and Sierra station. Lima station is the intersection of three lines, the Green, Red and Yellow lines, while all other transfer stations are at the intersection of two lines.

task with them. The training depended on which interface they were to use first (the SOI or the OAI). The training task involved exploring two stations and mapping out a journey between them. Typical interactions included finding the stations in the Index, finding out what lines they are on, using filters and/or the Legend to find out where the lines intersect and using the main overview and/or the Legend to get an idea of approximate journey times.

Tasks 1 & 2: Participants then carried out the first two meeting up tasks (see Appendices A.6 and A.7).

Description task: To break up the flow of the experiment and to ease the transition from the one interface to the other, I conducted a semi-structured interview about what participants had picked up about the public transportation system described by the application. Section A.8 contains the questions I asked. The interview took the form of role-playing in an attempt to bring out more casual discussion. The technique was based on teach-back methods as described in van der Veer & del Carmen Puerta Melguizo (2003).

Training: At this point of the experiment, participants changed from either the SOI to the OAI, or vice versa. If moving to the SOI, I explained to participants that they would no longer have access to any of the overview portions of the application and thus filtering would no longer be available. If moving to the OAI, I showed the participants the same sample task, showing them how to use the filters.

Tasks 3 & 4: Participants then carried out the second set of meeting up tasks (see Sections A.9 and A.10), using the interface that they did not use for Tasks 1 and 2.

Exit interview: At the conclusion of the experiment, I asked the participants to share with me what they understood about the public transportation system compared to the previous description task. The form of this interview is in Section A.11.

Payment: After the interview, participants were paid £10 for their time.

The flow of tasks

Regardless of which interface participants used first, they completed the same four tasks, shown in Table 4.4, in the same order. This is because each task builds upon the previous task and slowly introduces participants to the entire transportation system. Each task contains two lines and one transfer station that the participant has seen before. The idea is to simulate how a person might get to know an unfamiliar

Task	Starting stations					
1	Ida	(Yellow)	November	(Green)	Zeppelin	(Purple)
2	Bravo	(Purple)	Uniform	(Blue)	Washington	(Green)
3	Harry	(Red)	X-Ray	(Blue)	Zara	(Purple)
4	Freddy	(Red)	Marcel	(Yellow)	Uncle	(Blue)

Table 4.4: **The four meeting up tasks.** The first column denotes the task number and the following three show the starting stations for each task with their line name in parentheses. Participants were only given station names and no line information, as shown in Sections A.6 through A.10.

	Transfer stations									Lines				
	Baker	Charlie	Delta	Fiji	Lima	Oscar	Raoul	Sierra		Blue	Green	Purple	Red	Yellow
Task 1			x	x	x						x	x		x
Task 2		x	x				x			x	x	x		
Task 3	x	x				x				x		x	x	
Task 4					x	x		x		x			x	x
	1	2	2	1	2	2	1	1		3	2	3	2	2

Table 4.5: **The flow of tasks.** This matrix shows what transfer stations and lines each task can use. Note that these transfer stations are *useful* transfer stations; participants won't necessarily discover all of them. Each task follows from the previous one in that certain pieces of information are repeated. Each task shares two lines and a transfer station with the previous task. For example, Task 2 uses the Blue, Green and Purple lines; the Green and Purple lines were also used in Task 1. The most useful transfer stations in Task 2 are Charlie, Delta and Raoul; Task 1 also uses Delta station. The last row of the matrix shows how often these transfer stations and lines are used in the whole experiment.

transportation system: the same lines and stations crop up several times with the aim of making it possible for participants to build up a mental model of the transportation system. A random task order would have emphasised the task rather than knowledge of the transportation system, which is part of what was being investigated. For example, in Task 1, the relevant lines are the Green, Purple and Yellow lines and the most useful transfer stations are Fiji, Delta and Lima. In Task 2, the Green and Purple lines occur again and Delta is once more a useful transfer station. Table 4.5 shows the various lines and stations that a participant might encounter in a task. Note that while I can be assured that participants encounter the lines as designed, I can be less assured that they discovered all the transfer stations through the course of the task. This means that the repetitions of the lines is more apparent to the participants than the flow of useful transfer stations.

4.3.3 How the application supports the goals of the experiment

As I stated in Section 4.3.1, the goals of the experiment were to draw out the difference in interaction depending on the presence of the overviews and to identify any task efficiency that incurred from the

use of overviews. There are three ways the application supports those goals. First, it is an intentionally complex application and requires good understanding from the participants in order to use it. Second, the public transportation system described by the application is relatively busy and is not easy to fully grasp in one sitting. Third, the meeting up task itself is of higher complexity than a standard route-finding task and encourages more advanced usage of the application. Bringing these three types of complexity together results in the participants having to adapt to the constraints; the goal was to bring out greater variance in how they completed the tasks, depending on whether or not the overviews are present.

A complex application

The transportation system application uses a simple construct: a tree. This is not what makes the application complex. However the various layers in the application can make it difficult for participants to remember where they are in the application. Differentiating between the Legend and the Index is simplified by each side of the application using a different voice.

Remembering their location in the application is more difficult. For example, in the Index there is a list of stations on the first sub-level (Level 3 in Figure 4.1) and again on the last level (Level 7 in Figure 4.1). The main difference is that on Level 3, the list is alphabetical and contains all the stations in the system while on Level 7, the list is per line and is in travel order.

A medium-sized transportation system

The public transportation system represented in the application is a medium-sized one: it contains 73 stations on five different lines. In comparison to the London Underground (260 stations), it is smaller than the network that services Zone 1. It is comparable the Bielefeld, Germany[4], Washington, D.C., USA[5] and Stockholm, Sweden[6] Metro systems. It is not one that the participants could easily become comfortably familiar with over the course of the experiment. The idea is that the system would be representative of a real world one and, in order to accomplish the tasks, the participants would have to use the full extent of the application to explore the transportation system. Though the tasks have been designed to flow one to the other and slowly increase the participants' knowledge of the transportation system, that is more from the desire to create a comfort zone rather than from any expectation that the participants become experts on the transportation system.

---

[4]62 stations, http://metro.webseb.de/
[5]86 stations, http://www.wmata.com/
[6]100 stations, http://sl.se/

An intricate route-finding task

In a standard route-finding task, the goal is to find a route between two given stations. In the meeting up task, participants are only given half the picture. They are given starting stations but not destination stations. The task is to come up with the appropriate destination. This gives the participants an added challenge as it is more involved and will take more time. It is also representative of a real-world task that often takes a combination of fore-knowledge and cross-referencing of information.[7]

4.4   Conclusion

Understanding how people use overviews and their benefit to an application is a complex problem. On one hand, there is the overview and what can be understood from it. This is the question of whether the overview can convey the right information. Many of the auditory overviews analysed in Chapter 3 focus on this aspect of overviews. On the other hand, there is the role of an overview in context. That is to say, how does the overview affect interaction when combined with its detailed data, as opposed to when the overview is used in isolation. The aim of the application and its two interfaces described herein is to look at that more complicated question.

The metro map application provides an arena to study the two interfaces using a task which is sufficiently complex to facilitate an examination of the differences in usage with and without the presence of an overview. It is also a challenging application for an auditory interface as it represents a graph with many interconnections and as such could benefit from an overview to help guide the user. The next step is to investigate with an experiment how users react both to the presence of an overview and its use during interaction with the detailed information to which it refers. The meeting up task is a layered information seeking task that relies on participants coming to an understanding of how the metro system fits together. Paired with the description task, the experiment studies how the participants navigating patterns and understanding of the metro system differ depending on access to the overviews. The characteristics that emerged from Chapter 3 provide a framework for predicting the expected results of the experiment: improved understanding of the metro system, and more efficient information seeking and navigation patterns. Chapter 5 describes in detail how the overview changed the way participants solved the tasks.

---

[7]An   example   of   this   issue: http://goo.gl/ecs0j9   and   a   meeting   point   app:
http://jean-robert.github.io/metro.html

Chapter 5

Analysing usage of an overview-augmented auditory application

In participants' usage of the auditory public transportation system application, they showed a different pattern of behaviour depending on which interface they were using. The most significant change in their usage patterns was in how they went about finding information: certain parts of the interface showed more frequent usage than others. In the following sections, I describe these results and discuss what they mean in terms of the usefulness of an auditory overview.

5.1   About the participants

Twenty-two participants took part in the experiment, completing a total of 81 tasks. As mentioned in Section 4.3.2, two hours were allocated for each experiment. On seven occasions, the experiment took longer and a participant did not complete the fourth task. As a result, I have an unequal number of data points: 39 SOI tasks versus 42 OAI tasks. However, removing all experiments where not all four tests are completed has no affect on the statistical significance of the results and thus they were left in.

The twenty-two participants were recruited from the School of Electronic Engineering & Computer Science at Queen Mary University of London and all except one were students in those faculties. In the pre-questionnaire (see Section A.4), I collected basic data on each participant: their age, their sex, the number of years of computer experience they have, whether they have musical experience (what instruments, how many years they have played it and what they consider their level expertise), any
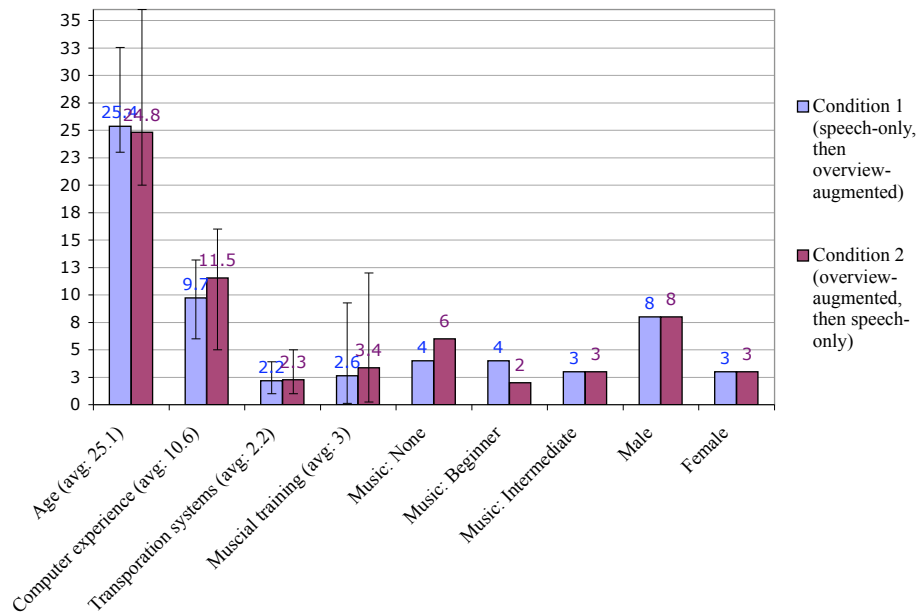
Figure 5.1: **The participants of the auditory transportation system application experiment.** The x-axis shows the categories of demographics collected. The majority of participants were in their mid-20's and highly computer-literate. On average, participants were familiar with slightly more than two different public transportation systems. There was a varying degree of musical experience: the average was around three years experience however, almost half had no experience and almost a quarter had ten or more years experience with a musical instrument. Just over a quarter of the participants were female, and the majority of them were in their mid-20s.

hearing problems they have, the number of different public transportation systems they are familiar with and what, if any, other research studies they have participated in the last year. Figure 5.1 summarises the data collected.

I recruited sighted participants with a computing background for several reasons: first, I was looking for a high level of comfort with computer usage; second, I had easy access to large pool of students; and third, as stated in Chapter 2, auditory applications are equally important to sighted users and therefore using them for testing is beneficial.

5.1.1 Age, sex and computer experience

The age, sex and computer experience information was collected to give a general impression of the participants involved. Computer experience was particularly important as it can give an indication of how comfortable a user is with the computer. Auditory interfaces are often intimidating to first time users and I wanted participants to feel as comfortable as possible using the interface.

The average age of the participants was 25 and the average number of years of computer experience was 10.6. Most participants had therefore been using computers since they were approximately 15.

Additionally, as they were students in Computer Science or Electronic Engineering, they were advanced computer users. This facility with computers meant that asking them to interact with the computer in an unfamiliar manner was not as far-fetched as someone who was a novice computer user.

The sex of the participants was not particularly crucial to the experiment but still important to keep balanced; 27% of my participants were female which is representative of the pool of computing students within the School.

5.1.2   Experience with public transportation systems

As the application displays a public transportation system, familiarity with one or more metro systems is useful. Knowing more than one system makes it easier for the participants to conceive of another system. Participants ranged from having experience with only the London Underground to knowing four other public transportation systems. Half the participants in each condition were familiar with only one or two systems while the other half were more seasoned travellers.

5.1.3   Musical experience

There is contradictory evidence regarding whether musical ability has an effect on performance in auditory interaction (Brewster, 1994; Edwards *et al.*, 2000). As the transportation system application has many rhythmic qualities and the main overview has a distinct pattern with a set number of pitches, participants were evenly allocated to groups based on musical training and years of experience to mitigate any potential effect.

5.2   The form of the data

The main data source for my analysis came from logging the participants' actions. I also collected various information from the interviews. Video-recordings were used to confirm notes taken during the interviews and any irregularities in the logs but where not transcribed for analysis. The logs of each task recorded every keystroke: this included what key was pressed to perform an action, the result of that action and a time-stamp of that action. Table 5.1 shows an excerpt of a typical log. An entire log from the experiment is included in Appendix C. The logs and the complete data table are available for download at `http://www.eecs.qmul.ac.uk/~lou/nickersonThesis/logs.xml` and `http://www.eecs.qmul.ac.uk/~lou/nickersonThesis/rawData.csv` respectively.

```
<nav time="152.753" key="right" location="0 0 2 ">
  —— the menu —— the Legend —— purple
</nav>
<nav time="154.772" key="down" location="0 0 2 0 ">
  —— the menu —— the Legend —— purple —— stations on the purple line
</nav>
<nav time="157.53" key="down" location="0 0 2 0 0 ">
  —— the menu —— the Legend —— purple —— stations on the purple line
  —— Nicolas station
</nav>
<nav time="162.291" key="left" location="0 0 2 0 15 ">
  —— the menu —— the Legend —— purple —— stations on the purple line
  —— Ancona station
</nav>
```

Table 5.1: **Excerpt of a log.** The `time` attribute is the time stamp of the action in seconds; the `key` attribute tells me what keyboard key the participants pressed in order to get to that location; the `location` attribute is the tree-path to the current node in the tree; the text value is the human-readable path into that particular node in the tree.

5.2.1 Typical task strategies

Typical task strategies varied. However, in order to understand the results, it is important to understand how a task might be answered. During training, I introduced participants to the following task solving process to get them started:

- Finding the starting stations in the Index,

- finding out what lines they are on,

- using the Legend/filters to find out where on the lines those stations are,

- identifying transfer stations through the filters or by exploring stations to the right and to the left of the starting stations in the Legend until they find one that matches the line of one of the other stations,

- moving back and forth between the various steps above, confirming information to reach an answer.

Before attempting their first task, I told participants that the goal is to come up with mutually convenient journeys for three people coming from the three starting stations. All the participants knew was the names of the stations and they must find out where in the transportation system the stations are, how to interconnect between them and decide upon a meeting station which gives approximately equal journeys to each person. A popular choice was to have people meet at a transfer station, resulting in two people

travelling directly to the meeting place and the third transferring to get there. Figure 5.2 shows an example answer to Task 2.

(a) Bravo

(b) Uniform

(c) Washington
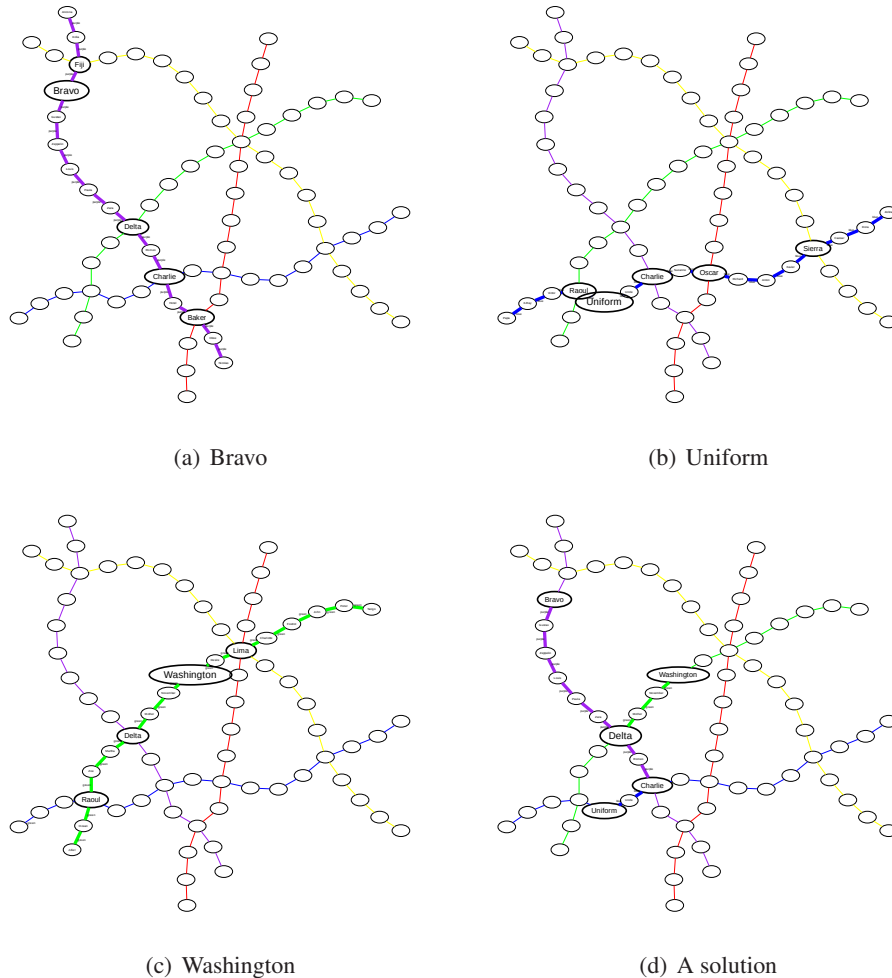
(d) A solution

Figure 5.2: **A solution to Task 2.** Figures (a) through (c) show the starting stations for Task 2: Bravo station, Uniform station and Washington station, respectively. These figures also show the various transfer stations on their respective lines. To solve the task, a participant would have to discover at least some of the transfer stations emphasised in each of these figures. Subfigure (d) shows a possible solution where the people at Bravo and Washington travel directly to Delta and the one at Uniform transfers at Charlie and then carries on to Delta.

5.3 Statistical methods

The analysis begins with basic task performance: solution correctness, completion time and number of moves within the application. This was to assess in very broad terms how the two interfaces compare in terms of efficiency and effectiveness. Also, the task times and scores were measures of how well the overviews gave a sense of the entire metro system and efficiently guided the participants to perti-

nent information (see predictions 2 and 3 in Section 4.3.1). To complete the analysis of participants understanding of the metro system, I evaluated the results of the description tasks.

To assess the participants' use of the menu system, I analysed where they went for information. This involved looking at moves made and time spent in the **Legend**, **Index** and at each level in the menu system.

The majority of the analysis is quantitative and compares measures of the **SOI** with that of the **OAI**. As each participant used both interfaces, though not necessarily in the same order but always performed the tasks in the same order, I needed to take into account the effect of the task on the analysis. This is a standard cross-over trial as described in Senn (2002). The results were analysed using a two-way ANOVA where possible or via Monte Carlo simulation. Where I used the ANOVA to analyse results, I report the F-value, the degrees of freedom of the interface (1), the degrees of freedom of the error and the p value, as specified in the APA style for reporting statistics. Where I needed the Monte Carlo simulation, I report simply the p value.

5.3.1 Rationale behind the statistical tests

One way the data could have been analysed is by pair-wise comparison of the two interfaces within each task by an independent-sample t-test. That is, comparing user behaviour using the **OAI** to that of the **SOI**. This type of analysis depends on all tasks and all participants being equivalent. The factor being analysed would have been the interface used. In the case of my experiment, one could argue that all the tasks were the same and only the three starting stations of the **meeting up task** varied. However, usually, to mitigate any unforeseen differences, the task order is varied. As task order was set – the tasks were designed to build upon each other – I could not pool the data for all the tasks and I would have been limited to comparing task 1 using the **OAI** to task 1 using the **SOI**, etc. In order to generalise about usage of the **OAI** versus **SOI**, I needed to take the task and the participant into account in the analysis.

The two-way ANOVA can take both task and condition into account and produce statistics that most strongly allow me to discuss differences between the two interfaces.

5.3.2 Techniques used

In my analysis, I looked at four major areas: task performance, **Legend** vs. **Index** usage, usage at each level of the menu and usage at each level in the **Legend** and **Index**. The total number of tests run was 61 of which 44 had significant differences. Table 5.2 shows the count of significant results across the

major areas. Appendix E contains all the ANOVA, adjusted mean and difference of mean tables for each measure.

The two-way ANOVA is a parametric test and thus the data needs to be normally distributed for the results to be valid. I ran the test for each measure and then graphed the normality of the residuals. 21 measures of the data turned out to be normal (within 98%) and statistically significant.

Where the residuals are not normal, a standard solution is to take the log of the data. For the remaining measures that could not be measured by a straight ANOVA test, I used log base 10 to run the ANOVA test. The results were then checked again for normality. This approach was valid and showed significance for 6 of the remaining measures.

For the measures where neither the straight data nor the log were normal but the ANOVA returned a p-value < 0.05, I ran a permutation test as described in Good (1997). Here, the data labels are scrambled and the two-way ANOVA is run on each possible permutation of the data. With 22 experiments, there are 22! possible permutations for the data. This is an incredibly large number and it would be very difficult to run them all. Thus, I used Monte Carlo Simulation (MCS) random sampling to generate 10,000 permutations of the data. The theory behind permutation tests and MCS is that if you run enough permutations and your result from the real data is ranked sufficiently highly amongst all the permutations, then the result is not explained by random chance. However, the data labels are not scrambled indiscriminately. Task order is not affected and the results from a single participant are kept together. For example, participant N could have used the speech-only interface for Tasks 1 & 2 and the overview-augmented one for Tasks 3 & 4; in the permutation, the labels either stay the same or are flipped to show that Tasks 1 & 2 were performed with the overview-augmented interface and Tasks 3 & 4 were performed with the speech-only interface. The F-values are then ranked and the p-value is the rank of the real data over the total number of permutations (i.e. 10,000). Of the 19 measures analysed with permutations, 17 showed statistical significance.

Over-testing

Given that I looked at so many measures, there is a risk that the results are chance as opposed to being actually significant. Over-testing is most likely when there are isolated measures that show significance. Confidence in conclusions about the findings also drops when statistics are chosen to support a desired outcome. The p-value reflects its likelihood of being chance and in most cases that is less than 1%. The level of confidence is strengthened where several statistics point in the same direction. For example, the

	Total	Significant	Straight	Log	MCS
Task performance	4	2		2	
Legend vs. Index usage	11	9	6	1	2
Level usage	16	15	8	3	4
Level usage in the Legend	15	9	4		5
Level usage in the Index	15	9	3		6
All	61	44	21	6	17

Table 5.2: **Categories of metrics and their rate of significance.** This table shows how many metrics were analysed. Straight inicates that the residuals were normal enough that I used the unadulted data when using the ANOVA when testing for significance. Log indicates that the log was used instead. MCS indicates that I used a permutation test with random sampling to validate the non-normal data tested with the ANOVA.

test for the usage of the Legend vs. the Index and the usage of the levels are almost all significant across the board. On the other hand, when comparing level usages in the Legend or Index, the results are not as strong so the conclusions drawn from those statistics are less trustworthy. The following sections report on each of the areas of analysis in detail as well as address the confidence in the results.

5.4   The results

5.4.1   Understanding of the metro system

Task performance

Task performance gives an indication of how the interfaces compare in terms of efficiency and effectiveness. The aim of the experiment is to look at whether the presence of an overview improves understanding of detailed information and how each interface is used differently. The basic metrics such as task completion time, number of moves, time per move and score show understanding of the metro system as well as how equivalent the interfaces are. Based on this low-level analysis, the two interfaces are not equivalent. They are equally effective but not equally efficient: tasks completed with the SOI are faster than those completed with the OAI. Furthermore, these statistics fail to show the expected effect of the overviews in terms of a deeper understanding of the metro system. Table 5.3 provides details of the statistics concerning task performance.

Task scores   Where the two interfaces are equivalent is in the quality of their answers. The criteria for a good score were efficient and effective journeys. Statistically, there is no difference between the scores. On average, tasks received a score of around 70% for each task (avg 69.2%). The addition of overviews does not give an advantage over the SOI. This means the interfaces were equally good at

Measure	Test data	F-value	*p*-value	Adjusted mean overview	Adjusted mean speech	Difference of means	SE of difference
Total time	log	7.37	0.01	3.06	2.95	-0.11	0.04
Num moves							not sig
Move/time	log	40.21	0.01	0.67	0.54	-0.13	0.02
Score							not sig

Table 5.3: **Task performance results.** The first column indicates what data was used to run the test; a blank entry indicates the raw data as opposed to the log or a MCS of the raw data. The next two columns show the statistics of the finding. The next two show the mean that has been adjusted to take into account task differences) for the OAI and the SOI. The final two columns indicate the difference between the two means and its asociated error. The negative number in the difference column indicates that the overview total time and time per move are larger than the speech ones. Both significant results required using the log of the data to ensure normality. The other two metrics showed no difference.

imparting information. This is logical as they both have the same underlying details as exposed through the speech system. It also means that the overviews did not present information in a way that makes it easier to answer the task but equally they were usable enough that they did not detract from finding the necessary information.

There were a small number of tasks that received a score of zero. This was not because the participants failed to complete the task but rather they specified impossible journeys. These impossible journeys were for the most part caused by problems understanding the text-to-speech. A discussion of these tasks is in Appendix D.

Calculating the scores The first step in calcuating the score is determining the distances between the starting points and the meeting station. The distance between two stations, d, is the number of stations between the two. d is then further adjusted to account for any transfers in the journey. Each transfer is counted as two further stations: number of legs, l, minus 1, multiplied by 2.

$$d = n + 2(l-1) \tag{5.1}$$

This factors in the amount of time it would take to perform that transfer. This is based on the assumption that whenever transferring from one line to another in the real world, additional time is taken to physically transfer between the lines and then the wait for another train to arrive.

The score for each task (S) is calculated using the distances from each starting station to the meeting place (d_{1-3}) by adding the average of the shortest and longest distances to the standard deviation of the three distances.

$$S = mean(min(d_1, d_2, d_3), max(d_1, d_2, d_3)) + stdev(d_1, d_2, d_3) \tag{5.2}$$

The first part of Equation 5.2 checks the range of distances and favours shorter journeys, penalising longer ones. The second part evaluates how equal the journeys are and rewards equal-length journeys, penalising unfairly different lengths. The combination of the two thus checks the length and evenness of the three journeys. Since both criteria were given equal weight in the instructions to the participants, neither is given more importance in the calcuation of the score.[1]

$$S_{final} = 100 \frac{S_{opt}}{S_{task}} \tag{5.3}$$

The final score (S_{final}) is the ratio of the optimal score for the task (S_{opt}) to the task score (S_{task}), scaled to give a number between 0 and 100. The optimal journey was determined by hand to match the criteria set out for the tasks. This ensures that the value of S_{opt} is always smaller or equal to S_{task}. The larger the task score gets, the smaller the final score will be. Scaling the ratio between the two by 100 provides a larger number so as to be easier and more familiar to discuss.

Task efficiency   Task efficiency is essentially the amount of time it takes to complete a task. Another part of this is how many moves (keystrokes) a participant makes. The number of moves is equivalent between the two interfaces. Where the interfaces diverge is on time per task and time per move.

Moves per task   The number of moves per task starts to give a better idea of what users do while performing their tasks. Each participant took on average 377 moves per task. OAI tasks are less consistent in the number of moves but the difference is not great enough to make the result significant. The implication of this is that the speech application drives the efficiency in investigating the detailed information and the overview, at least in this case, does not provide any improvement.

Task completion time   The first thing to note about task completion times is that they are long. For comparison, informal tests indicated that, visually, the task, given the same circumstances, takes about three and a half minutes. Pilots tests of the OAI took between 7 and 15 minutes to complete tasks. As such, completion times of 20 minutes, such as with the OAI were unexpected. The main difference between the pilots and the experiment was the formality of the setting: casually asking an acquaintance

---

[1]It should be noted that giving the first part of the equation a higher weighting results in unequal journeys getting a higher score. On the other hand, giving the second part a higher weighting makes no difference in a ranking of scores and has little effect on the number.

to run through an experiment is different than the pressure felt in a laboratory setting. This implies that a certain improvement in performance can be expected from approaching the tasks in a more relaxed manner. I would also expect some improvement as users become more familiar with the application. Regardless, the auditory application is unlikely to reach the performance of a visual map which speaks more about how we should use auditory applications than their inefficiency. An auditory application that is slower than a visual one might be the right tool in a particular context, such as an eyes-free one.

There is a lot of variability in task completion time, with both the shortest and longest times coming from the OAI. SOI tasks are more efficient with an average completion time roughly 25% less than OAI tasks (987 seconds vs. 1329 seconds, $F_{1,76} = 7.37, p < 0.01$).

Time per move Unsurprisingly, given the difference in completion times, move times are also longer with the OAI (2.5s/move vs. 3.8s/move, $F_{1,76} = 40.21, p < 0.01$). This result might be expected given that all station and line elements in the OAI contained an overview, requiring more listening, and participants would need to pause to listen to the main overview.

Interaction with the main overview filters

The results above show that overviews do not make the meeting up tasks easier nor faster to complete. One important question is whether participants used the overviews at all. There was a possibility that they just ignored the sounds and relied only on the speech in the interface. Participant usage of filters belies this. It is clear that participants made use of the main overview and were regularly manipulating it. What is less clear is the meaning behind the filter usage. However, their usage implies that the presence of the overviews affected how they completed their tasks. The statistics discussed below are shown in Table 5.4.

The functions a user can perform with the filters are setting them to focus on an individual line or station, removing filters so they no longer hear information about a line or station and clearing all filters to reset the main overview. On average, participants used these functions 16 times per task. With an average completion time for tasks using the OAI of 1331 seconds, this equates to participants performing a filter action every 83 seconds.

Furthermore, participants were setting several filters at once and used a variety of unique filters. On average, participants set 6.43 unique filters per task, with an average upper limit of 4.5 filters set at a time. Participants tended to clear all filters 1.81 times per task. With regards to usage of station and line filters, the interaction is evenly split; in a task, participants would use station or line filters about seven

times. On average, this works out to about 5.5 filters being set and about 1.6 being turned off.

The picture that emerges here is that participants listened to and manipulated the overview with the filters. This is also implied by the longer task completion time for OAI tasks.

	Average
Interactions with filters	16.07
Unique filters used	6.43
Most filters used concurrently	4.59
Clearing all filters	1.81
Interactions with station filters	7.05
Station filters turned on	5.55
Station filters turned off	1.50
Interactions with line filters	7.21
Line filters turned on	5.50
Line filters turned off	1.71

Table 5.4: **Filter usage in the OAI.** This table shows the average number of uses of filters in a task. The top four rows show general usage or interaction with the filters while the bottom six break down interaction by type of filter and state of the filter (setting or unsetting).

Description task

The completion times and scores reported above contradict the predicted results (see Section 4.3.1) with regards to getting a stronger mental model of the transportation system. Hearing the scope and structure of the system was expected to have led to better scores. Equally, the overviews' guiding aspect should have led participants more directly to the information they needed to complete the task. The final piece to look at is the description task. Do the overviews change the participants' understanding of the transportation system? I interviewed participants between Tasks 2 and 3 – as they switched interfaces – and again after Task 4. The purpose of the interviews was to draw out what they had learnt about the transportation system described by the application. The interviews after the first two tasks would set the basis for information gleaned from each interface. The interviews at the end of the experiment could identify the increase of knowledge after using the other interface. I questioned them on the number of lines in the system, their knowledge of transfer stations, the size of the system and the configuration of the system.

Participants generally knew that the system had five lines and were able to identify between one and five transfer stations. They were equally (in)accurate in identifying which lines went through the named transfer stations. Most participants agreed that the system was medium-sized – the London Underground was used as a comparison – and had no real conception about how it was laid out. The general impression was that the system was well-connected. Each participant was able to add and/or

correct a few details after Task 4 but there was no consistency in what details were gleaned after Tasks 3 and 4 nor any correlation to which interface was used for each portion of the experiment.

The results of the interviews were analysed by manually cross-comparing what the participants reported. There was no discernable difference between the knowledge of participants during the two description tasks. It was unlikely that a more in-depth analysis involving transcripts and coding of the videos would find any different so this line of analysis was abandonned. During the interviews, the participants did make some remarks about how they used the overviews and their preferences for one over the other. These are summarised in Appendix B. There were not sufficient comments for them to be taken as representative nor to report a significant preference for either interface; it seems instead that some participants took to the overviews quickly and were able to take advantage of the information encoded into the overviews, while others had difficulty adapting to the use of non-speech sound.

I had expected that the spatial relations provided by the overviews would lead to a better understanding of the transportation system. This was not the case. With the SOI, participants had to search manually for interconnections, while in the OAI, the main overview provided clues to potential routes through filters. It seems that the manual process used in the SOI allowed participants to recall interconnection information in the same manner as the main overview and filters did in the OAI. With the limited time that the participants spent exploring the transportation system, neither interface provided an advantage in terms of grasping the scope and structure of the metro system.

5.4.2 Where participants navigate

So far, the results of the experiment do not align with the hypotheses formed in Section 4.3.1 regarding understanding of the metro system. The other hypothesis dealt with how the overviews drive exploration of the detailed data. This section discusses where participants are going for information. The first navigation patterns I looked at was usage of the Legend versus the Index and the second was the depth and breadth of their usage of the application.

Index and Legend usage

As explained in Section 4.2.2, the Index provides a station-based view into the public transportation system while the Legend provides a line-based view. When participants wish to look up something about an unknown station, they can look it up in the Index; when they wish to explore a line, they can look it up in the Legend. Usage of these two sides of the application provides an insight into how the

Measure	Test data	F-value	*p*-value	Adjusted mean overview	Adjusted mean speech	Difference of means	SE of difference
Visits to Index	log	5.91	0.05	0.45	0.36	-0.09	0.04
Visits to Legend							
Number of switches							
Time in the Index		33.42	0.01	62%	30%	-33%	6%
TPT Index	MCS		0.01	43%	22%	-21%	6%
Time in the Legend		33.53	0.01	38%	70%	33%	6%
TPT Legend		44.89	0.01	20%	58%	37%	5%
Moves in the Index		27.88	0.01	65%	36%	-30%	6%
MPT Index	MCS		0.01	45%	27%	-18%	6%
Moves in the Legend		27.88	0.01	34%	64%	30%	6%
MPT Legend		37.17	0.01	19%	53%	34%	5%

Table 5.5: **Legend and Index usage results.** This table is organised in the same manner as Table 5.3. MPT and TPT are shorthand for moves/time per turn. Tasks completed with overviews tend to favour the Index, while those completed without overviews favour the Legend.

presence of overviews affects the participants' task solving strategies.

Changes of context between the Legend and the Index　Table 5.5 describes the various aspects of the Index and Legend that I analysed. The first aspect was how often participants used either side of the menu system. Looking at the count of times participants go to the Index or Legend and the number of times they switch, the only metric that returned a significant difference is the number of visits to the Index. Given that the three metrics should reflect each other, I am not confident there is any real difference between the two interfaces.

Usage of the Index　The Index gets used much more heavily by participants using the OAI. They spend more time there (ANOVA: $F_{1,76} = 33.42, p < 0.01$) and make more moves there (ANOVA: $F_{1,76} = 27.88, p < 0.01$). The same holds true for each visit to the Index: they make more moves per visit (MCS: $p < 0.01$) and spend more time per visit (MCS: $p < 0.01$). Given that the time spent and moves made are all significant, the confidence in these statistics is high.

Usage of the Legend　Participants using the SOI use the Legend more often than those using the OAI one. This finding shows up in all interactions with the Legend: SOI tasks spend about 33% more time in the Legend (ANOVA: $F_{1,76} = 33.53, p < 0.01$). Equally, SOI tasks make more moves in the Legend (ANOVA: $F_{1,76} = 27.88, p < 0.01$). Each visit to the Legend is far more thorough as well: participants spend more time (ANOVA: $F_{1,76} = 44.89, p < 0.01$) and make more moves per visit in a SOI task (ANOVA: $F_{1,76} = 37.17, p < 0.01$). Again, all the measures support each other so the confidence in these statistics is high.

Measure	Test data	F-value	*p*-value	Adjusted mean overview	Adjusted mean speech	Difference of means	SE of difference
Num unique nodes	log	13.51	0.01	2.09	2.19	0.10	0.03
Level 3 time		10.47	0.01	25%	19%	-6%	2%
Level 3 moves		35.53	0.01	44%	26%	-17%	3%
Level 3 unique							
Level 4 time		65.02	0.01	25%	9%	-15%	2%
Level 4 moves	MCS		0.01	15%	9%	-5%	1%
Level 4 unique	log	11.42	0.01	1.23	1.08	-0.15	0.04
Level 5 time		71.78	0.01	24%	46%	22%	3%
Level 5 moves		29.69	0.01	24%	40%	15%	3%
Level 5 unique		9.12	0.01	33.5	48.3	14.8	4.9
Level 6 time	MCS		0.01	16%	9%	-7%	2%
Level 6 moves		5.54	0.05	6%	9%	2%	1%
Level 6 unique		6.23	0.05	10.9	15.2	4.3	1.7
Level 7 time	MCS		0.05	5%	12%	7%	2%
Level 7 moves	MCS		0.01	5%	12%	7%	2%
Level 7 unique	log	14.08	0.01	0.7	1.2	0.5	0.1

Table 5.6: **Usage results of the levels.** Each condition used the various levels of the application menu differently. OAI tasks focused more on Levels 3 and 4 while SOI tasks focused on the lower 3 levels.

What this implies is that without the overviews, participants rely much more heavily on the information that the Legend provides, i.e. the organisation of the transportation system is via the lines. The Legend is one of the two places that participants can go to hear the stations in order on each line. As shown in Figure 4.1, level 5 of the Legend and level 7 of the Index lists stations in order per line. The Legend is also the only place that participants can easily explore neighbouring stations as there is no equivalent in the Index.

Usage of the seven levels of the application

The difference in usage of the Legend and the Index indicates that participants need different information depending on which interface they are using. This is further backed up by analysing the usage of the seven different levels in the application.

Knowing what levels participants navigate to helps establish what questions they are trying to answer. Each level in the interface has a particular usage. Level 3, as shown in Figure 4.1, is the first level of enquiry and answers the query "tell me more about [. . . station/line]." Going deeper into the application represents sub-queries where the participant has found something out and wishes to know more about it.

For example, a participant looks for *Alpha station*: this interaction takes them to Level 3 in the Index. To find out more about Alpha station, the participant presses <down> twice to level 5 to find out

what line it is on. Discovering that Alpha station is on the Yellow line, the participant could go down two more levels to Level 7 to find out what stations are on the Yellow line or where Alpha station is on the line. Each further delving deeper into the application is to satisfy the desire to know more.

So where and how deep into the application are the participants going to solve their tasks? To address this question, I examined navigation at each level in terms of number of unique nodes visited, amount of time spent and number of moves. The unique nodes visited indicate the breadth of the exploration at each level while the number of moves and time spent indicate how focused the participant is on the information provided by that level.

With OAI tasks, the focus is on Levels 3 and 4. This implies that they search for stations and lines, set filters and then move on to the next line or station of interest. They use Level 6 as well but only in terms of time, not number of nodes or moves. This implies that they are not exploring at that level but rather navigating quite directly to some information; I speculate that they are going there to set a filter then listen to the effect. If this is the case, then this is representative of the overviews acting as a guide and directing the participants to information of use. See Table 5.6 for the specifics on the statistics.

With SOI tasks, the participants focus mostly on Level 5. This is indicative of them following up on information as they seek. For example, if they are exploring a line, they drill further down to explore the stations on that line. They also focus on Levels 6 and 7. The navigation pattern here implies a pattern of data exploration: finding information leads to needing further information. It is a pattern of questions leading to sub-questions.

Another thing to note is that participants using the OAI visit 20% fewer unique nodes overall (ANOVA of the log: $F_{1,76} = 13.51, p < 0.01$). Given that both conditions use the same number of moves per task (see Section 5.4.1), this implies that participants using the OAI go to the same nodes repeatedly.

Usage of the 7 levels in the Legend and Index

The analysis of the logs shows that participants using the SOI prefer to explore the metro system via the lines (i.e. in the Legend) and tend to do nested data exploration (i.e. search for information and then drill deeper into the menu hierarchy). Participants using the OAI, on the other hand, tend to use the Index (i.e. navigate by station) and interact more shallowly in the menu hiearchy. However, since the roles of the two sides of the menu are different, what the participants are doing is still unclear. Isolating the activity at various levels in either the Legend or the Index shows in more detail not only where participants are

Measure	Test data	F-value	*p*-value	Adjusted mean overview	Adjusted mean speech	Difference of means	SE of difference
Legend L3 time	MCS						
Legend L3 moves							
Legend L3 unique	MCS		0.01	3.3	4.8	1.4	0.3
Legend L4 time	MCS						
Legend L4 moves							
Legend L4 unique							
Legend L5 time		62.37	0.01	15%	41%	26%	3%
Legend L5 moves		28.71	0.01	17%	36%	19%	3%
Legend L5 unique		11.92	0.01	26.6	41.7	15.1	4.4
Legend L6 time							
Legend L6 moves	MCS		0.05	3%	8%	5%	1%
Legend L6 unique		16.96	0.01	5.2	12.9	7.7	1.9
Legend L7 time	MCS		0.05	2%	8%	6%	1%
Legend L7 moves	MCS		0.05	2%	8%	6%	1%
Legend L7 unique	MCS		0.05	3.1	13	9.9	1.8

Table 5.7: **Legend usage at each level.** SOI tasks use the Legend more actively however, there is similar usage at Levels 3 and 4.

navigating but indicates what information they are seeking. Tables 5.7 and 5.8 show the statistics for time spent, moves made and unique nodes visited at each level in the Index and Legend.

While there were differences accross the board when looking at Levels and the Legend/Index, there are some similarities: Levels 3 and 4 of the Legend and Level 7 of the Index show similar activity. While the OAI tasks use Levels 3 and 4 more overall (see Table 5.6), it turns out this is only in the Index. Similarly, SOI tasks use Level 5 more overall, but this is only in the Legend; in the Index, OAI tasks us it more.

In order to get a better idea of what this all means, I generated a heat map of where each condition
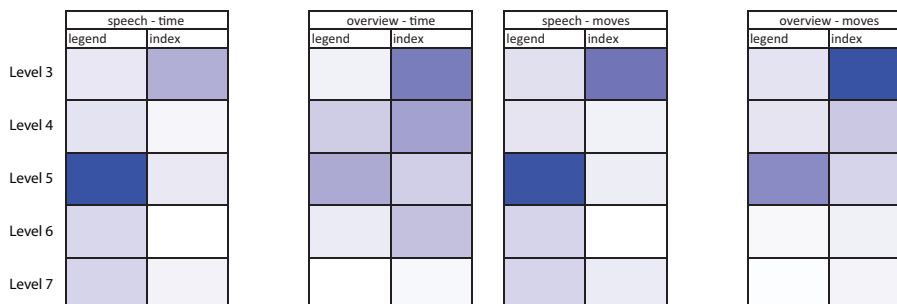
Figure 5.3: **Activity at each level in the Index and Legend.** This figure shows the relative usage of nodes by the participants, split up by level. The colour of the left two is based on the relative amount of time spent in a part of the application. The right two are coloured by the relative number of moves. The time spent and moves made with the SOI show the focus of the participants on the stations in the system, while the interaction patterns in the OAI are much more varied.

Measure	Test data	F-value	*p*-value	Adjusted mean overview	Adjusted mean speech	Difference of means	SE of difference
Index L3 time		13.68	0.01	21%	14%	-7%	2%
Index L3 moves		31.64	0.01	4%	2%	-2%	3%
Index L3 unique							
Index L4 time		55.51	0.01	16%	3%	-12%	2%
Index L4 moves	MCS		0.01	9%	4%	-5%	1%
Index L4 unique							
Index L5 time	MCS		0.05	9%	5%	-4%	1%
Index L5 moves	MCS		0.05	7%	5%	-3%	1%
Index L5 unique							
Index L6 time	MCS		0.01	11%	1%	-9%	2%
Index L6 moves	MCS		0.01	4%	1%	-2%	1%
Index L6 unique	MCS		0.01	5.8	2.4	-3.4	0.8
Index L7 time							
Index L7 moves							
Index L7 unique							

Table 5.8: **Index usage results per level.** he number of unique nodes visited is the same at every level however OAI tasks use the Index more than OAI tasks at all levels except for Level 7.

focuses their activity. The heat map shows several things of interest: first, in the SOI tasks, where the participants move and where they spend time is the same. Second, participants primarily get their information from Level 5 in the Legend. This is where the menu lists what stations are on each line. The participants still use more of the lower levels than when using the OAI. The pattern of interaction that emerges is that 50% of time and moves are focused on the stations, whether that be in the Legend at Level 5 or in the Index at Level 3. They appear to follow a logical strategy: find the stations in Index, find what lines they are on, find the lines in the Legend, explore what stations are on the line, drill down further to identify transfer stations.

On the other hand, OAI tasks so not have such a clear pattern. Where the participants move is not always where they spend time. Their move patterns look similar to the SOI pattern in that the densest spots are Index Level 3 and Legend Level 5. Except they make half as many moves in the Legend Level 5 (ANOVA: $F_{1,76} = 28.71, p < 0.01$) and twice as many moves in Index Level 3 (ANOVA: $F_{1,76} = 31.64, p < 0.01$). Note also that they visit as many unique nodes in the Index Level 3 so they are making repeat visits to the stations. The OAI time heat map looks unlike the others. Participants seem to spend time more evenly across the entire application instead of focusing on specific places. This implies that they are accessing a broader set of information and exploring the system more fully. In several places (such as Index Level 6) they barely make any moves, yet spent time there. It is difficult to determine what the interactional pattern is however, I speculate that it is the overviews that are sending

the participants to investigate various lines and stations. They move more in the Index Level 3 as they investigate stations hihglighted in the overview. They then navigate more broadly elsewhere in comparison with when using the SOI to investigate pieces of information as directed by overview and dwell once they confirm information or set a filter. If this is the case, then there is a degree of interaction efficiency (though no time efficiency), of being guided by the overview and encouraged to explore.

5.5 Discussion

The goal of the experiment described in this chapter was to look at how interaction with a user interface is affected by the presence of an overview. Section 4.3.1 hypothesised that overviews would have two effects: the first is that participants would have a greater understanding of the metro system that the application describes. Additionally, I predicted that task completion time and task accuracy would improve as a result of participants' knowledge of the scope and structure of the metro. The second effect would result in a more direct and directed navigation pattern which would result in a faster completion time. Also, the navigation pattern would be broader as the overviews would encourage the participants to explore. In general, there is little evidence that the overviews improved efficiency in this manner. However, there was a different infomation-seeking pattern that hints that the latter hypothesis is valid.

5.5.1   Understanding of the metro system

The understanding of the metro was primarily measured by assessing the quality of the answers and querying the participants knowledge about the structure and scope of the metro. Neither of these metrics showed that the OAI had any advantage over the SOI or vice versa. Since the participants did not develop a better model of how the metro interconnected, the faster completion times did not follow.

The long completion times were, in fact, a surprise. Completion times were expected to be on the order of seven minutes but with the overviews, they were easily three times that. The culprit was move times: completing the tasks with the OAI did not take more moves than with the SOI, just more time. One likely explanation for the longer move times is that participants receive a lot more information as they interact with the application. Every time they navigate to a station or a line, they hear an overview; every time they make a change to a filter, they must listen to the main overview. This means that they are constantly receiving information to listen to and process, while as when using the SOI, participants only receive information when they request it by navigating through the application. This aspect of using

overviews must be taken into account when designing an auditory application.

5.5.2 Navigating the application menu

As seen in Chapter 3, Overviews act as guides and encourage exploration. It follows that the navigation pattern should be more efficient. There is a definite difference in how the participants approach the task when they have access to the overviews. What needs to be measured is whether there is any *interaction-efficiency* since it is clear there there is no time-efficiency. What is meant by interaction-efficient is that overviews could make it possible for users to get more information with less effort/movement: the listening that is required by overviews increased how much time it took participants to use the application, however the interaction was more straightforward and thus more efficient. The SOI interface saw participants methodically exploring stations organised by lines. The majority of their time was spent in the Legend on Level 3 but they also performed nested info-seeking tasks driving them deeper into the menu hierarchy. For example, one of their goals was likely to be "where are the transfer stations?" In order to find relevant transfer stations, they need to know what lines to look at. In order to know what lines ot look at, they need to investigate the starting stations. This work flow explains the patterns of activity in Figure 5.3.

On the other hand, there are indications that, with the OAI, the participants were doing more with their moves. For example, participants visit the same number of nodes in the Index on Level 3 but are there for longer and make more moves. This implies that they are exploring more stations than their SOI counterparts. This could be because the overviews tell them about more stations and they go to investigate them. They use more areas in the menu system for their task yet fewer unique nodes. The implication is that they know where to go but need to navigate to get the information so more of their interaciton is in getting to information to confirm it rather than discovering it. Once they get to the information, they then stop to listen. This type of interaction would explain the differing patterns of time spent and number of moves in in Figure 5.3.

That said, there are a number of other possibilities. At the other end of the spectrum, participants could have been confused by what they heard in the overview and felt the need to confirm several times. They could have also been having problems of too high an information load. This too would result in an erratic pattern of moves and time. However, if this were the case, the participants would have been stopping and listening at all levels rather than primarily at levels where they do not move much. Their pattern is much more of one of directed exploration.

5.5.3 Conclusions

The overviews in this experiment did not contribute to a discernable better level of knowledge about the metro system. Nor did they improve the quality of the answers to the meeting up tasks. In this sense, the first hypothesis, which deal with understanding of the structurem the scope and the salient features of the metro, is rejected. On the other hand, the presence of overviews had an impact on how the participants navigated. The reasons for the differing pattern are not clear however, the evidence implies that the overviews led participants to pertinent information. With a speech-only application, users hear only detailed information and need to piece things together themselves. For simple tasks, this is not so much of an issue; with advanced interaction, where auditory applications strive to be as rich and efficient as graphical ones, the ability to represent compound information such as in an overview is crucial. Overviews can provide a shorthand for information that would otherwise have to be searched for manually. The end result would be a more interaction-efficient auditory application if not necessarily a more time-efficient one.

Chapter 6

Conclusion

This thesis has two primary contributions: the first is an in-depth analysis of overviews as an interface artefact; the second is an evaluation of how the presence of an overview affects interaction in an auditory interface. These are relevant to a broad range of fields, but most specifically interaction design, usability and auditory interfaces. In this chapter, I discuss future work and the ramifications of my research, describing the findings and their limitations.

6.1 Contributions and discussion

6.1.1 The nature of overviews

Research into overviews often erroneously assumes that we know what the term means. My analysis of overviews in Chapter 3 is novel in that it aims to operationalise[1] the term *overview*. Understanding what overviews are allows myself and other researchers to address the area from a clearer standpoint.

My approach was systematic: review what others say about overviews in order to generalise about them. Then, conduct an independent survey of several commonly found and varied overviews to come up with a list of characteristics that all the overviews have in common. Then, look back at the initial research reviewed through the lens of the findings of the role of auditory overviews. The end result was a list of six characteristics that should be present in order to call something an overview. This finding has two implications: first, it allows us to identify overviews even if they are not called so and to eliminate

---

[1] a term from social sciences that means to concretise a concept so that it is clearly delineated what it does and does not comprise

false overviews; second, it permits us to understand what impact an overview may have as part of an interface.

Reviewing research on overviews

The research on overviews (Section 3.1.1) covers a range of interfaces and applications. This provided me a broad base for examining what an overview is. However, there appears to be no significant research specifically on overviews. Rather, the main topic of the research reviewed is the interface and its usability as opposed to defining and examining the characteristics of overviews. This means that the body of work that I took as my basis contains a watered-down description of how the authors qualified overviews: they explained only as much as they needed in order to address the main topic of their publications. Additionally, I only have the authors' descriptions of an overview and could not assess the overviews themselves in the manner that I did the visual overviews. Despite this lack of information, the descriptions seem valid as they align with the findings of the survey that followed. In other words, the themes in the literature are corroborated by the findings of the survey.

In general, the researchers who have used overviews seem to have a good idea of what an overview is. There is a sense that it should provide an organised view of the detailed information. It is the exceptions to this that make my research necessary. In some work, an "overview" seems synonymous with metadata. In other words, the "overview" consists of a list of unconnected facts. This is akin to, for example, summarising a document with its number of words, pages, chapters and its file size – informative, but useless for understanding anything about the document without reading it. Pérez-Quiñones *et al.* is an example of this sort of approach where they provide a list of options the user can perform. A table of contents, on the other hand, displays richer information which provides some understanding of the underlying document.

Another approach I have criticised is thumbnails as overviews. In some cases, one can rely on the layout of the detailed information to express overview-type information. Where this falls apart is when the detailed information is not laid out in such a way that miniaturises well. For example, a map made smaller can be useless, however a map that understands what is a major versus minor landmark can prioritise what to display at various magnifications. An example of using thumbnails as overviews is Hornbæk & Frøkjær (2001) where zooming is the primary overview mechanism. However, the power of thumbnails cannot be ignored. They are useful, for example, in slide sorter interfaces as seen in

presentation applications such as Microsoft PowerPoint.[2]   On their own, thumbnails provide a weak overview but as part of a more sophisticated overview, they are an expressive interface artefact.

Surveying visual overviews

The above examples of inconsistencies show that we need a firmer definition of overviews. The six characteristics allow us to be more discriminating about what is and is not an overview. A clearer understanding of the role of an overview should play can also help determine if an overview is needed or desired. Shneiderman (1996) talks about the crucial role of overviews in info-seeking. The domain of info-seeking is important to his assertion: it implies a large data set. A small one might not require an overview. And asking oneself if all or most of the overview characteristics are necessary will expose that.

The six characteristics are a product of a systematic analysis. I looked at a diverse set of visual overviews to extract their commonalities. Each overview was broken down into its components and their function, and then analysed from the point of view of a typical usage. This allowed me to study the overviews from several different angles. However, the analysis was not exhaustive and it represents one researcher's point of view. To strengthen the approach, I also relied on others' assertions about those overviews. As a result, I feel confident in the validity of the findings.

There are three major, shown here in bold, and three minor overview characteristics:

- **showing scope**,

- **quickly guiding**,

- **showing context**,

- describing the structure,

- encouraging exploration and

- providing a historical snapshot.

The major ones appeared in all the overviews I analysed and the minor ones in most. This reinforces an important aspect of overviews: that they are flexible and their form should depend on their intended usage. Some overviews are well enough understood that they can be automatically generated. This is the case with tables of contents. However, this is not to say that their format is rigid. Different applications generate them differently. One application might make heavy use of indentation, another of font size. A table of contents for a collection of works by different authors may include author names

---

[2]http://office.microsoft.com/en-us/powerpoint/

for each piece while that would be superfluous for a table of contents for a document with a consistent set of authors. In other words, an overview needs to be appropriate to its intended usage. The variation in the six characteristics support this. This is backed up by work by Tversky where she addresses maps as overviews notes the following characteristics: "they omit information, they regularize, they use inconsistent scale and perspective, and they exaggerate, fantasize, and carry messages" (pg.1). This speaks greatly to maps needing to be realistic only to the point that they are useful, allowing them to become social devices. Where visualisations and overviews of data sets try to be as accurate as possible, other forms of overviews distort to convey a message Tversky (2000). In a nutshell, overviews are flexible.

Treating the six characteristics as absolute contravenes the intent of an overview. They need to be applied as necessary rather than by rote. This can be seen in the variety of combinations of the characteristics both in the overviews I analysed (see Table 3.5) and in the overviews I reviewed from the literature (see Table 3.18).

Looking at the six characteristics as a collective, there are several themes that run through them. The most obvious is how overviews represent the depth and breadth of the detailed data. Most of the descriptive characteristics (see Table 3.5) contribute to this theme. In essence, these characteristics answer the question "what is this all about?" The characteristics are how to achieve the theme. Describing the scope most readily maps to describing the breadth; clarifying the structure answers many questions about the depth of the detailed data; context also further details how topics are addressed and the historical snapshot characteristic can describe the growth or evolution of the data. All of these together help users – whether as creators or consumers of the content – assess what they can expect to find in the detailed data.

Another theme that runs through the characteristics is exposing interrelationships. This theme is most readily supported by structure and context. Interrelationships are not just about salient features; they are also about showing connections that are harder to perceive when the user is interacting with the detailed data. This is the same notion as not being able to see the forest for the trees.

A final theme is about interaction with the detailed data. I mentioned in Chapter 4 how overviews are at their most powerful when used in conjunction with the detailed data. In order to encourage interaction, the overview must facilitate it. While the other two themes are about showing what information is there and providing new insight into it, this theme is about delving into the detailed data to find out more.

This is where, in particular, quickly guiding and encouraging exploration come into play. Providing easy access to the detail takes the overview from the passive "show" to the more active "investigate." For example, a line graph showing a timeline might highlight that something of interest occurred in a particular month. The graph gives the user the knowledge of which month and s/he can then go on to search out the raw data from that month to see the particulars. The other aspect of interaction is encouraging exploration. This may be through exposing an interesting interrelationship, prompting the user to delve into the detail; it may also be from questions that arise from delving into the detail. In this second case, the user might flip between the detail and the overview to gain better understanding of the information provided by both.

In summary, the six characteristics that emerged from my research in Chapter 3 describe in more detail some major themes about overviews. The diversity of overviews surveyed and reviewed show that overviews need to be adaptable as opposed to rigid, meaning that there is a great deal of flexibility in creating them. The major themes that are common to the topic of overviews are describing the breadth and depth of the detailed data, showing interrelationships and facilitating interaction. The six characteristics map out more specifically how to satisfy these themes.

6.1.2 Overview-guided interaction

In Chapters 4 and 5, I moved from the theory of overviews to how they operate in practice. Most research to date focuses on successful usage of an overview and its usability. Instead, I chose to focus on what difference an overview makes. The main reason for this is to understand how people use an overview in practice and how the overview can affect the design of an application. While this thesis is not about design, there are implications from this thesis that can be applied to that area. Another focus of this work is auditory interfaces. Many researchers who work in auditory overviews believe, like I do, that overviews can move auditory interfaces forward to be richer and more useful/usable, as discussed in Sections 2.2.3 and 3.1.1.

My approach was straight-forward: take an auditory application and compare usage depending on the presence of an overview. The auditory interface was a fairly typical one that made use of speech as its primary output. It was non-trivial however and I required participants to solve a composite task in order to pull out sufficiently rich interaction for me to study. The results were analysed quantitatively and showed a distinct difference in interaction patterns when overviews are present but no improvement in terms of efficiency and effectiveness.

The main novelty in the experiment was the task I asked participants to perform. Rather than asking them to find individual answers contained in the detailed information, the meeting up task required them to search out various pieces of information and reason about them. As a result, I had extensive logs to analyse. The results of the experiment were quite clear: access to overviews changes how a participant navigates the application and looks for information. The quantity of data collected (see Appendix C for an example of the data collected from one of the 81 tasks completed by participants) reinforces the statistical confidence in these findings.

The other novelty of my research was comparing a plain auditory interface with an overview-augmented one. Prior work on auditory overviews primarily focused on evaluating the performance of the overview (see Sections 2.2.3 and 3.1.1). I was specifically interested in assessing how an overview changes interaction in order to strengthen the argument for including them. With a greater understanding of what happens when they are not present, I hope to mature the field of auditory overviews with the end goal that overviews be a standard part of auditory interfaces.

One clear finding of my experiment is that in terms of effectiveness the interfaces are equivalent. The number of participant moves within the application were statistically the same. Also, the journeys specified for the meeting up tasks were the same. The major difference was in task completion time: The overviews made the tasks slower. In essence, the presence of the overviews in the Overview-Augmented Interface (OAI) made things slower but otherwise performed miuch like the Speech-Only Interface (SOI). In other words, the addition of an overview does not translate into better performance.

On the other hand, I was able to show that overviews can have an impact on interactional patterns. There were differences in which view participants primarily used (index versus legend) and where they concentrated their activities. The presence of overviews resulted in more diverse usage and less extensive usage of the legend. This implies a pattern of listening and confirming information rather than gathering of information that occurs with the SOI.

The clear difference between the two conditions in the experiment implies that applications should be designed differently depending on whether overviews are included in the interface. The difficulty then is knowing how to alter or design an application to accommodate overviews. One issue is that regardless of what information is encoded in an auditory overview, one is unsure what is being perceived by the user. This is not unique to auditory overviews but it is more acute due to the immaturity of the field and the difficulty of using, for example, a talk-aloud protocol as the participants would be simultaneously

talking and listening (Chandrashekar *et al.*, 2006; Strain *et al.*, 2007). Psychoacoustics can tell us what is possible for the user to perceive and work like Bregman (1994) tells us a lot about how humans separate sounds into streams. However, there is less research on what happens when we apply these principles to an interface and the user is confronted with using the information provided in the audio. From my experiment, there is evidence that, given an overview, users will listen and make use of it. Further studies of this nature will more strongly affirm how users react to the presence of an overview – and what they find acceptable, – increasing researchers and designers confidence in predicting the effects of overviews.

Clearer expectations of the effect of overviews would provide a better idea of how to design auditory applications that include overviews. Similar to how a table of contents can affect the creation of a document, I would argue that the inclusion of a overview should affect how the detailed data is presented in an application. Applications could be optimised to use the overview to its best effect. The different interaction patterns in my experiment show that parts of the application had different levels of usage and these patterns can be used to iteratively optimise the interface. My experiment tested only one interface and one can assume that the effects of an overview would be different with other interfaces. In that sense, to develop a general guide to using overviews, overviews need to be tested in conjunction with the detailed data as opposed to in isolation. I would further argue that this sort of user testing should become an integral part of the development of auditory interfaces. A foreseeable outcome would be more streamlined and effective auditory applications which may lead to better interfaces for visually-impaired users and opening up auditory applications to the wider market of support for eyes-free tasks.

One of the most successful classes of applications that make use of audio are GPS navigation systems such as TomTom[3] or Ovi Maps on Nokia phones[4]. Speech is the primary audio output and it provides the detailed instructions to the driver. The visual display provides overview-like information, taking the driver's eyes away from the road. Arguably, the visual display is best able to mimic what the driver will see ahead of them. However, it also highlights the current industry's reluctance to include more sophisticated audio in a mainstream application. Audio is seeing some success in specialist applications such as stock market monitoring (e.g. Childs, 2005; Worrall, 2009), seismology (e.g. Dombois, 2001; Joseph & Lodha, 2002) and medical applications (e.g. Hermann *et al.*, 2006). Their adoption can likely be attributed to the advantages that real-time audio monitoring can provide. The findings in this thesis indicate there is much room for improvements which can move auditory interfaces into the mainstream.

---

[3] http://www.tomtom.com/
[4] http://store.ovi.com/content/100493

6.2 Future work

In order to bring about more widespread usage of auditory overviews, let alone auditory interfaces, the work in this thesis must be refined and built upon.

6.2.1 On further defining overviews

I studied overviews in Chapter 3 in order to apply that knowledge to auditory overviews. The six characteristics that emerged were first steps to properly understanding overviews. The next steps are to build upon my findings through investigating wider examples of overviews and broader feedback on how overviews assist users.

There are several things that studying a wider range of overviews could accomplish. First, a broader survey could identify more subtle characteristics. Second, we could begin to identify types of overviews and create a classification based on type of overview and difference between modalities. I propose an analysis similar to the one I have already conducted which looks at typical usage and identifies the various components of the overview. A broad range of overviews should be selected across various application areas and modalities to allow the greatest diversity of characteristics to emerge. Which, in turn, will increase the overall understanding of how to apply overviews. While this can be seen as mostly an academic exercise, it would provide a firmer basis for the inclusion of overviews in computer interfaces. As we clarify their role, the advantage of including them in an interface should also become clearer, justifying their usage.

For the most part, users interact with overviews instinctively and without much thought. An in-depth study, in the form of unstructured interviews and focus groups, could draw out uses intended and unintended by overview designers – those involved in devising new overviews – and creators – those involved in the content of the overview. The resulting data could be used to affirm the characteristics of an overview and delve more specifically into various types of overviews. In essence, participants would reproduce the analysis I conducted in Section 3.1.2 to understand how users perceive overviews and put them to use. The group interviews or focus groups would provide a dynamic environment where each participant's contributions could trigger richer information.

In essence, the approaches described in the previous two paragraphs have the same goal: to draw out richer understanding and more fully explore the topic of overviews. The questions that would be addressed would be: are there more characteristics? Are there different types of overviews? Do the

characteristics depend on overview type or modality?

6.2.2 Investigating overview-augmented auditory interfaces

Getting back to auditory overviews, I have already discussed the fact that the findings presented here are the results of one interface in one experiment. The results were specific to the transportation map application; more generally, they indicate that overviews change the user experience. However, the results bring new questions to light: how would a different overview affect the usage of the interface? How do overviews work in other interfaces? What happens when an overview is used to guide the design process of an auditory application? To answer these questions, the work I have started here needs to be expanded and generalised by studying in more depth how overviews can support the transportation map application, the common effect overviews have and how we can capitalise on the impact of overviews on an interface.

In the case of the auditory transportation system described in Chapter 4, overviews affected the navigational patterns within the application. Would the effect be the same given a different overview? An experiment evaluating different overviews on top of the SOI would focus on what the most appropriate overview would be for the application. Specifically, each overview could be evaluated across several metrics: task efficiency, understanding of the transportation systems, mental load, etc. The findings would resolve how different overview approaches are received by users and develop guidelines for picking the appropriate mechanism for specific tasks. This is paramount to understanding how auditory overviews can be introduced into mainstream applications.

The next research question focuses on generalising across multiple uses of overviews. Will other auditory applications with overviews show the same results? Will the effect of overviews be of a different nature? In what cases do overviews not change user behaviour? This is about finding common effects of overviews. Several applications with and without overviews would need to be cross-compared to identify similarities and differences in how overviews impact their usage. With this knowledge, we would gain a better understanding of how and when to apply overviews to an interface. This approach necessitates reproducing others' overviews and incorporating them into full applications. Only Zhao *et al.* appear to have built their overview into an interface that included the detailed data. However, they did not study what happens when the overview was not present. Re-evaluating their experiment, and conducting more experiments like this, could identify a broader range of effects of overviews on interaction with the detailed data.

The last research question takes the research forward to a new level and builds upon the two areas of research detailed above: what is the effect of applying this knowledge to the design process? Ultimately, the goal is to create better and more auditory interfaces that are acceptable in the mainstream. Through this thesis, I posit that auditory overviews are an integral part of the solution. Each of the above-mentioned interfaces could be optimised and improved – be it the overview or the presentation of the detail – and subsequent evaluation of these interfaces would learn the effect of incorporating overviews into the design process. The hypothesised outcome is that this will allow for the development of better quality auditory applications that will be easier to use for sighted and visually-impaired users alike.

This three-pronged research agenda is in pursuit of a better understanding of auditory overviews and how they can evolve the field of auditory interfaces. The first step is understanding how different overviews support the usage of an auditory application and builds directly on the experiment described in Chapters 4 and 5. The second step is determining if similar effects emerge with different auditory applications and generalise the current findings. The third builds on the knowledge from the previous two steps and establishes how overviews can be an integral part of developing auditory applications.

6.3   Summary

The two main aims of this thesis are to operationalise the term *overview* and to analyse the effect of auditory overviews on an auditory interface. One motivation behind the work is to move auditory interfaces into mainstream usage. In this chapter, I have discussed the implications of my research findings and outlined a research agenda that will push both aims forward.

The analysis of visual overviews provided a basis for understanding what overviews are and the role they play in an interface. The characteristics that emerged help define their applicability in a variety of uses. The future work described seeks to further refine and generalise the findings by repeating and broadening the analysis. By deepening the knowledge in this area, we can further understand and classify the use of overviews and prevent the misapplication of the term. This goal will provide a platform for the application of overviews in other fields.

Furthering this work entails studying how overviews can contribute to the field of auditory interfaces. This firmer understanding will enable us to determine the appropriate overview for the task, find what effects of an overview are repeatable and determine whether integration of overview evaluation in the design process results in more acceptable auditory interfaces. Ultimately, the evaluation of an auditory

overview in practice showed that it can impact how users interact with an application.

Glossary

amplitude The instantaneous amplitude of an oscillating quantity is its value at any instant, while the peak amplitude is the maximum value that the quantity contains (International Community for Auditory Display, 2011). 25

direction The localisation of the sound source. We determine direction by the difference in time and intensity between each ear. 25

duration The length of a sound from attack to decline. 25

frequency For a sine wave the frequency is the number of periods occurring in one second. The unit is the cycles per second, or Hz. For a complex periodic sound the term 'repetition rate' is used to describe the number of periods per second (p.p.s.) (International Community for Auditory Display, 2011). 25

highlight Setting a filter on a station or a line in the main overview to make it more obvious. A station overview causes the sound of that station to be time-stretched and a line overview makes other lines drop out of the overview. 86, 87

index The list of station names in the auditory transportation system application. Allows users to look up stations and find out what lines they are on. 8, 82, 83, 85, 96, 98, 103, 105, 107, 112–119, 126

legend The list of lines in the auditory transportation system application. Allows users to look up lines and find out what stations are on them. 8, 82, 83, 85, 88, 96, 98, 103, 105, 107, 112–117, 119, 126

line overview A non-speech overview used in the auditory transportation system application that describes a single line: a single tone representing that line plays and delineates where the first and last stations are in terms of the whole system. The overview plays when a user navigates to a line node in the application. 84, 90

loudness This is the intensive attribute of an auditory sensation, in terms of which sounds may be ordered on a scale extending from quiet to loud (International Community for Auditory Display, 2011). 25

main overview A non-speech overview used in the auditory transportation system application that describes the entire system at once. The audio presents all the stations in the system by playing a short sounds representing the lines that go through them. Users can apply filters on lines and stations to draw out information of interest. This overview loops continually throughout usage of the application but fades into the background when the user is navigating so it does not interfere with the speech output. 84

meeting up task A complex route planning task used in assessing the auditory transportation system application where participants had to map out a convenient meeting place for three people given three start stations. 9, 92, 94, 96–99, 105, 110, 120, 126, 151

options level A level in the auditory transportation system application that lives below any station or line node. It allows users to explore the line/station and set filters on the main overview. 82, 86

phase The phase of a periodic waveform is the fractional part of a period through which the waveform has advanced, measured from some arbitrary point in time (International Community for Auditory Display, 2011). 25

pitch Pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale (International Community for Auditory Display, 2011). 25

simple station A station in the transportation system that exists on only one transportation line. 86

spearcon A short representation of speech where the speech is sped up to the point of incomprehensibility.. 26, 87

station overview A non-speech overview used in the auditory transportation system application that describes a single station: the tone(s) representing that station play(s) and delineates where in the systems that station is. It plays when a user navigates to a station node in the application. 84

timbre Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and have the same loudness and pitch are dissimilar. Put more simply, it relates to the quality of a sound (International Community for Auditory Display, 2011). 25

time-stretch A change in the duration of an audio event without affecting its pitch. In the auditory transportation system application, highlighted or filtered stations are time-stretched in the main overview. 87

transfer station A station in the transportation system that is on more than one line. 86

Acronyms

AISA Auditory Information-Seeking Actions. 40

ASA Auditory Scene Analysis. 25, 28

EEG Electroencephalogram. 16, 17, 30, 31

GUI Graphical User Interface. 20, 24, 29, 33–35

HCI Human-Computer Interaction. 15, 16

HDS High-Density Sonification. 42, 43

HRTF Head Related Transfer Function. 33

MBS Model-based sonification. 31, 32, 43

MCS Monte Carlo Simulation. 106–108, 113, 114, 116, 117, 195, 197, 198, 203–205, 208–210, 212–216

OAI Overview-Augmented Interface. 8, 10, 79, 84, 87, 90, 93, 94, 96, 100, 105, 107–119, 126, 151–153

SOI Speech-Only Interface. 6, 79, 81, 84, 86, 94, 96, 100, 105, 107, 108, 110, 112–119, 126, 129, 151–153

TTS Text-to-Speech. 28, 29

VI Visually-Impaired. 17–24, 33, 34, 36–39, 45, 46

Appendix A

Auditory public transportation system application experiment files

This appendix contains the forms used in the auditory public transportation system application experiment described in a Chapter 4. For clarity of display, the forms start on the next page.

A.1 Recruitment circular

Auditory metro map experiment

We're looking for people to participate in an experiment using an auditory application to navigate a metro system, like the London Underground. The experiment will take 2 hours and you will be paid £10 for your time.

If you are interested in participating, please contact **Louise Nickerson**:

Phone ⊔⊔⊔⊔⊔ ⊔⊔⊔ ⊔⊔⊔

Email lou/map@...

Further details

This research aims to improve the usability of audio-only applications. Audio-only computer interfaces often present a challenge because of the amount of information a user must remember. To study how complex information may be represented in audio, we have developed an auditory metro map application.

The purpose of the study is to evaluate how an auditory overview of the metro system enables you to perform route finding tasks. Please remember that it is the auditory representation that is being tested, not your memory or your ability to find the most expedient routes.

Procedure and tasks

You will perform a series of route-finding tasks and description tasks using two versions of the application: one version is speech only and the other is augmented by overviews. The route finding tasks involve mapping out journeys from various locations and the description tasks involve describing what you know about the metro system.

Notes

The subject under investigation is the auditory interface, not you - we are not assessing levels of expertise.

All your personal data is anonymised in the analysis of this study. When reporting on this study no information about you is revealed.

A.2 Information sheet

Info sheet - Auditory metro map

Thank you for participating in this study!

This research aims to improve the usability of audio-only applications. Audio-only computer interfaces often present a challenge because of the amount of information a user must remember. To study how complex information may be represented in audio, we have developed an auditory metro map application. The purpose of the study is to evaluate how an auditory overview of the metro system enables you to perform route finding tasks. Please remember that it is the auditory representation that is being tested, not your memory or your ability to find the most expedient routes.

Procedure and tasks

We will both record your experience on video as well as track your actions through the computer interface. Please let us know if you are uncomfortable with this.

There are 4 parts to this study:

1. We ask you to complete a short pre-questionnaire. This will probe for demographical information and aspects of your experience with computers and audio.

2. We will explain the application you will be using and run through some sample tasks.

3. You will perform 4 route-finding tasks and 2 description tasks using two versions of the application: one version is speech only and the other is augmented by overviews. The route finding tasks involve mapping out journeys from various locations and the description tasks involve describing what you know about the metro system.

4. Finally, we will interview you about your experience with the auditory application.

Notes

The subject under investigation is the auditory interface, not you - we are not assessing levels of expertise. All your personal data is anonymised in the analysis of this study. When reporting on this study no information about you is revealed.

A.3   Consent form

```
This form was filled in duplicate, one copy for the participant, one
copy for my records.  The text in italics was read out loud before
the participant signed.
```

Consent Form

Participant _____

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree to take part.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

I understand that if I decide at any time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately.

I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

I consent to the video of this experiment to be kept on file for future interaction studies. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

⊔ Please tick here if you agree to be contacted regarding future experiments conducted by the Department of Computer Science. If so please provide your email address:

_____

Participant's Statement

I _____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: Date:

Investigator's Statement

I _____ confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer.

Signed: Date:

A.4   Pre-questionnaire

Pre-questionnaire

Participant _____

Please ensure that you have read and completed the consent form before filling out this questionnaire.

Age	
Sex	
Number of years computer experience	
Musical experience	
Instruments	_____
Years played	_____
Level	Advanced / Intermediate / Beginner
Please list any hearing problems	
Have you participated in or are currently participating in any research studies in the last 12 months?	Yes/No (circle one)
Number of different metro systems you are familiar with	

A.5 Presentation for participants about the experiment

Metro map experiment

What is this experiment about?

- What are we assessing?
  - The impact of overviews in an auditory application on the user experience
- What is an auditory application?
  - An application that uses sound instead of visuals
- What is an overview?
  - An overview is a brief summary. Some examples:
    - The table of contents of a book
    - A blurb describing a museum exhibit
    - The directory tree in your file manager

The flow of the experiment

- What's involved? (this presentation)
- Signing of the consent form
  - Allow us to collect data on your usage of the application including video.
- Pre-questionnaire
  - Name, age, computer experience, experience with metro systems, etc.
- Training period
  - Journey planning tasks: how to get from A to B
- Experiment tasks
  - Meeting up tasks: where should people coming from A, B and C meet?
  - Description tasks: describe the metro system
  - 2 tasks using **speech only**, 2 tasks using **speech + overviews** (order will vary)
- Post-experiment interview
  - Thoughts about the experiment

So what happens now?

1. What kind of application are we testing?
   - A metro map application
2. What will I be doing?
   - Route-finding tasks
3. How do I use it?
   - Keyboard interaction
4. What will I hear?
   - Speech - stations and lines read out
   - Non-speech - beeps and bleeps that tell you about an action you've performed
   - Overviews - sounds that describe the metro system - you will learn how to interpret these as you proceed

---

The metro map application

- The map application consists of menu system. It has:
  - A legend: lists all the lines in the metro system
    - Links to stations on those lines
  - An index: lists all the stations in the metro system
    - Links to the lines through those stations
  - This is an imaginary metro systems
    - Lines have colour names
    - Stations have names derived from phonetetic alphabets (alpha, bravo, charlie, etc)
- You will use the application in two different modes
  - **Speech only**: the menu items are read to you
  - **Speech + Overviews**: the menu items are read to you and supplemented by
    - An overview of the entire system
      - This overview is customisable through filters on stations and lines
    - Overviews of individual lines
    - Overviews of individual stations

The tasks

- 4 Meeting up tasks: where should people coming from A, B and C meet?
  - You will be given 3 station names.
  - Your task is to plot routes to a central meeting place from those stations
  - The goal is to make the journeys as equal possible for each person
  - You should describe routes like this:
    - Russell Square: FEW stops to Leicester Square on the Piccadilly line, FEW stops on the Northern line to Waterloo
- Description tasks: describe the metro system
  - After every 2 tasks, you will be asked to give a description of the system to someone
  - You will be given a verbal prompt
  - Your task is to respond being as helpful as possible
- You will perform 2 meeting up tasks and a description task using the **speech only** application
- You will perform 2 meeting up tasks and a description task using the **speech + overviews** application
- Which application you use first will vary!
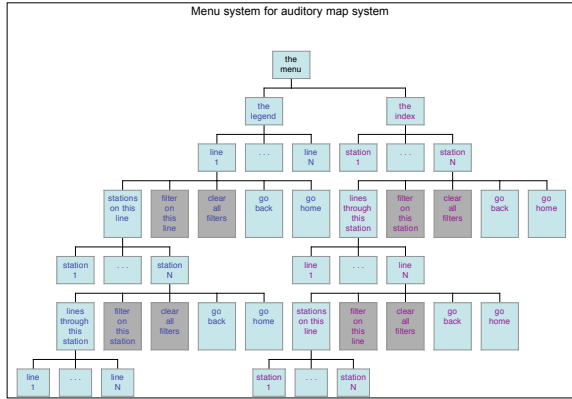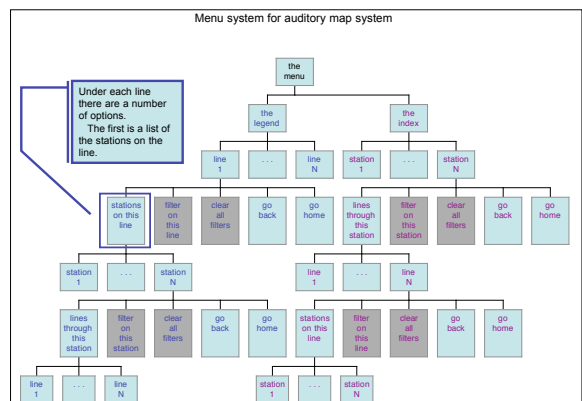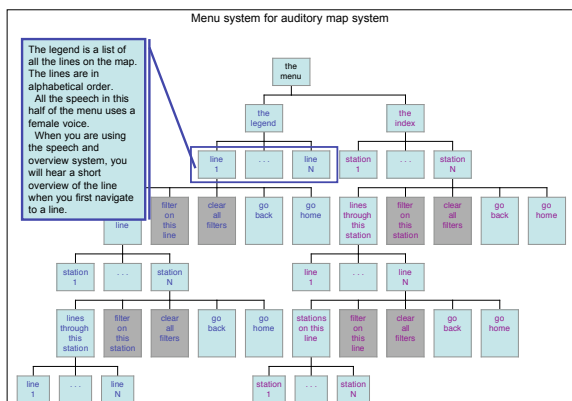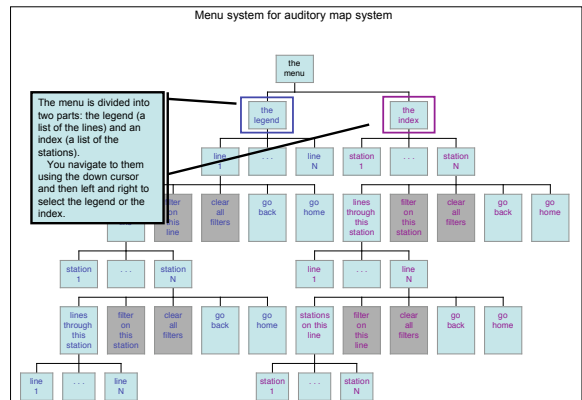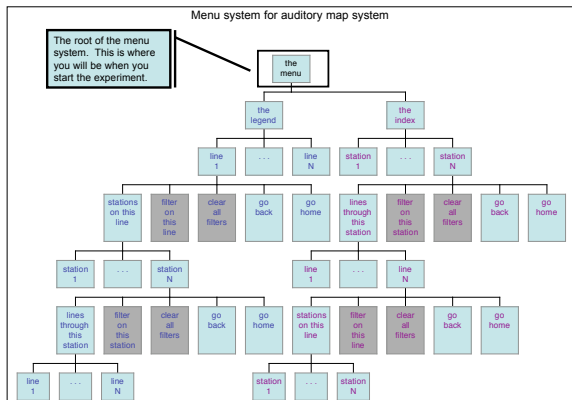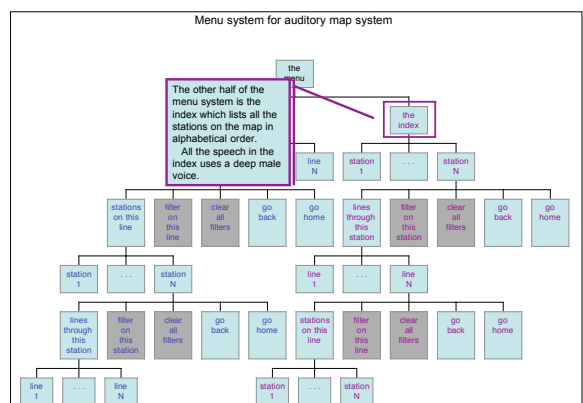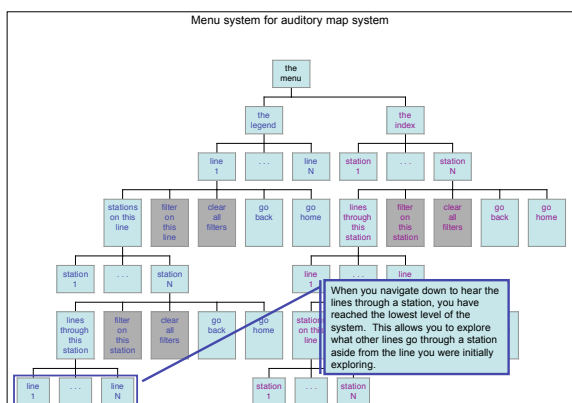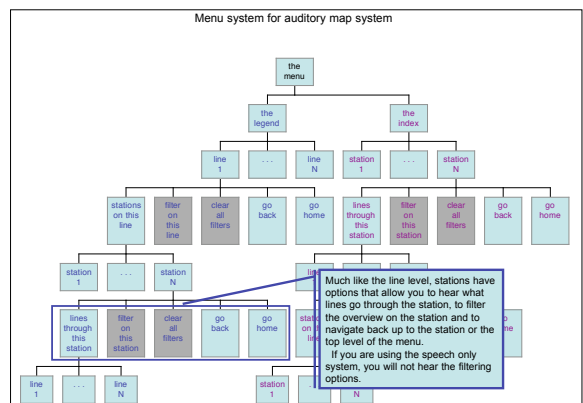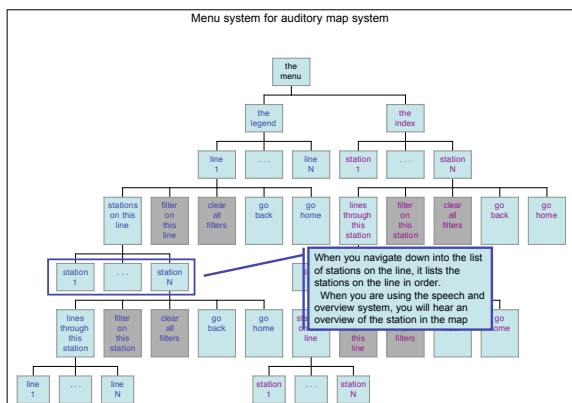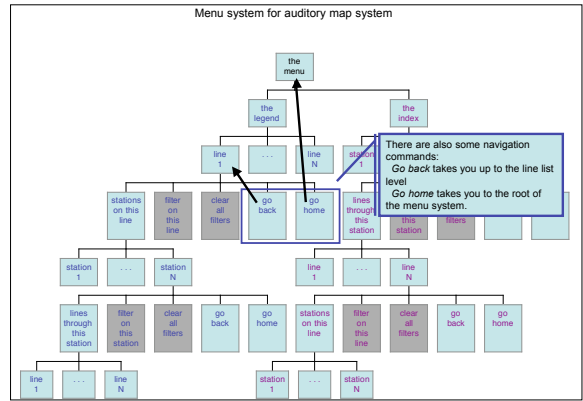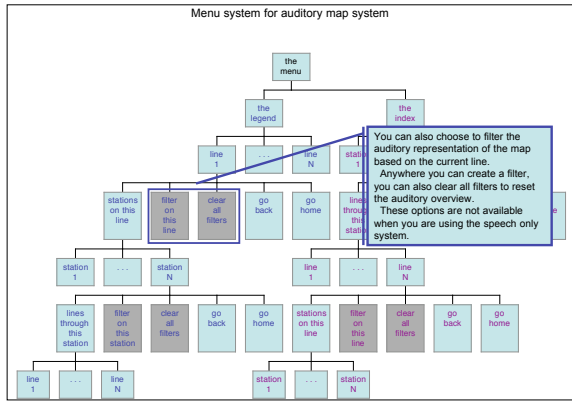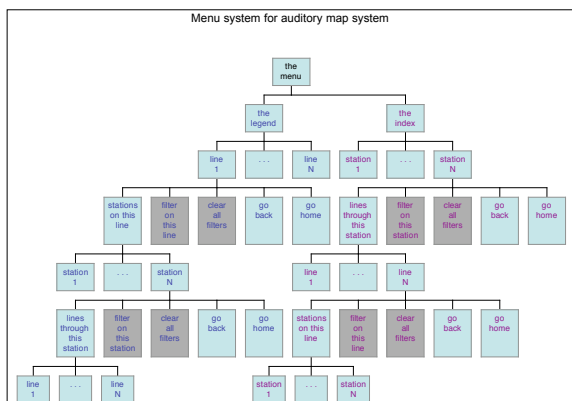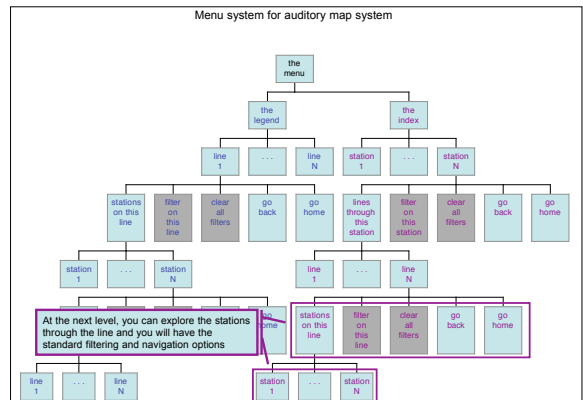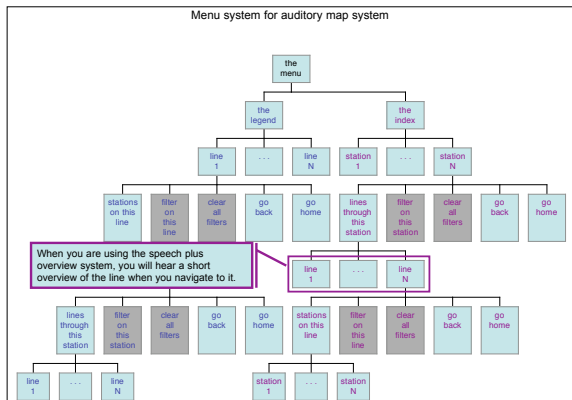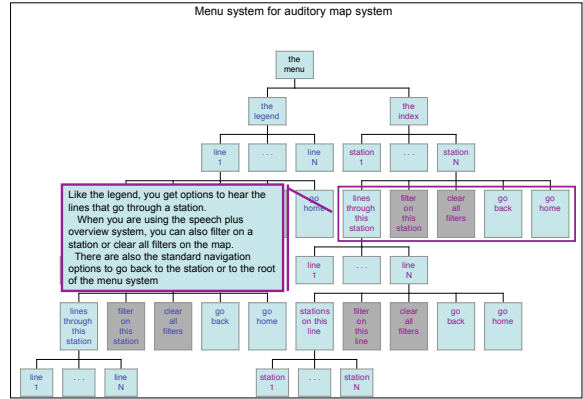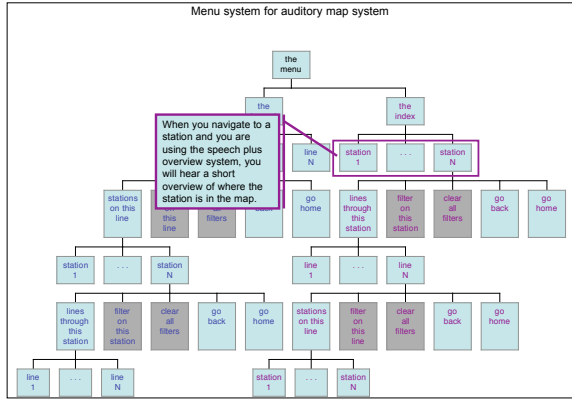
A.6   Meeting task 1

Participant _____

Determine a meeting place for three people coming from the following stations:

Ida — November — Zeppelin

Meeting place: _____

Route from **Ida**:

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Route from **November**:

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Route from **Zeppelin**:

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
      (circle one)   (line name)   (station name)

A.7   Meeting task 2

Participant _____

Determine a meeting place for three people coming from the following stations:

Bravo — Uniform — Washington

Meeting place: _____

Route from **Bravo**:

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Route from **Uniform**:

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Route from **Washington**:

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
  (circle one)   (line name)   (station name)

A.8 Description task

I would read the prompt after explaining that we would be role-playing. They were instructed to be as natural and helpful as possible. I would gradually work the four target areas below into the conversation.

Prompt

Hi! I'm Susan and I'm moving to Alphabet City next month. I know you just moved there yourself but I'm wondering if you can tell me about the metro system there? You know, just to get the lay of the land a bit. Where are the big hubs? Someone told me that Washington or Delta station would be a good place to live?

Hubs Identified	Lines in the system ....
Delta (green & purple)	red
Lima (green, red & yellow)	green
Raoul (blue & green)	blue
Charlie (blue & purple)	yellow
Fiji (yellow & purple)	purple
Sierra (blue & yellow)	5
Oscar (blue & red)	4
Baker (red & purple)	3
Size of system	layout of lines
small	bunched at the beginning
medium	don't know
large	spread out

A.9   Meeting task 3

Participant _____

Determine a meeting place for three people coming from the following stations:

Harry — X-Ray — Zara

Meeting place: _____

Route from **Harry**:

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Route from **X-Ray**:

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Route from **Zara**:

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

Many / Some / Few stops on the _____ line to _____ station
(circle one)   (line name)   (station name)

A.10 Meeting task 4

Participant _____

Determine a meeting place for three people coming from the following stations:

Freddy — Marcel — Uncle

Meeting place: _____

Route from **Freddy**:

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Route from **Marcel**:

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Route from **Uncle**:

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

Many / Some / Few stops on the _____ line to _____ station
‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(circle one)‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾(line name)‾‾‾‾‾‾‾‾‾‾‾‾(station name)

A.11 Description and exit interview

Prompt

```
I would summarise their previous responses, leaving them time to cor-
roborate or contradict what they said before.  It was entirely free
form.  Then I would ask them if they had any comments about their ex-
perience with the two versions of the application.  As this was not
a usability test, I steered them away from comments on problems they
had with the interface.
```

You said _____ last time. Can you tell me more now?

Appendix B

Participant feedback regarding interface preference

Chapter 4 describes an experiment investigating how the addition of an overview to a speech-based auditory map application affects its usage. I interviewed participants both mid-way through and at the end of the experiment to capture their understanding of the metro system. These results are reported in Section 5.4.1. Here, I address what participants said about their experience with the two interfaces.

Several participants made comments about their strategies for solving the meeting up tasks. The meaningful comments were sparse and cannot be considered representative of all the participants. However, they shed some insight into how they used the overviews. The comments mostly emerged during the second description task and fell into two general categories: preferring the clarity of the Speech-Only Interface (SOI) or preferring the added information in the Overview-Augmented Interface (OAI). Participants are denoted by P$n_{ssoo|ooss}$ which indicates their number and the order they did the tasks ($o =$ OAI and $s =$ SOI).

B.1 The clarity of the SOI

A few participants expressed the notion that while the overviews could provide additional information, they preferred the simplicity of the SOI. From their comments, it appears that they felt pressure to interpret the overview and as a result, the overview was a hindrance as opposed to an aid. This was expressed most clearly by P23$_{ssoo}$.

P23$_{ssoo}$ started with the SOI for tasks 1 & 2 and then moved on to the OAI. P23$_{ssoo}$ took as long on

task 3 as on tasks 1 & 2 combined. It was clear from observation that P23$_{ssoo}$ was struggling with the interface change. In the interview, P23$_{ssoo}$ talked about how the overviews were challenging: *"I tried to differentiate each sound but I could not [...] At first, I tried to use the sound but [....] it was easier not to."*

Another example is P2$_{ooss}$ who expressed frustration at not being able to hear more of the layout. P2$_{ooss}$ used the OAI first and commented in the first description task: *"I have no idea of layout; it's really confusing"* and again *"I can't draw the map."*

What's striking about these examples is the expectation that they should immediately be able to extract information from the overview with very little familiarity. As sighted users used to graphical interfaces, they set the bar for themselves high. They expect easy-to-use and intuitive interfaces and in an laboratory setting, they also want to perform well, despite experimenters explaining that it is the interface being tested and not them. In the post interview, P23$_{ssoo}$ asserts that now that all the tasks are done, P23$_{ssoo}$ has a better understanding of the overviews and would do much better. This speaks to P23$_{ssoo}$'s desire to perform well, plus the potential benefits of a longer settling-in period.

Familiarity and modality issues aside, this sentiment, which was not expressed unilaterally, does convey that the overview was complicated and hard for them to grasp. Listening to non-speech sound was difficult for some and they would have preferred to be able to turn the overview off. This is not unexpected since most people are only used to trivial sounds in computer interfaces. Introducing people to an interface which uses complex sounds manipulated in various of ways is bound to overwhelm some users. One solution is obviously let the user decide if s/he wants use the overview; the challenge then is to make the overview so useful or convenient so as to overcome the user's reticence much in the way Palm did in the use of the Graffiti handwriting recognition (MacKenzie & Zhang, 1997).

B.2   How the overviews augmented the application

More frequently, participants mentioned missing overviews when they switched interface. Several participants talked in their exit interview about how they used the overview and how they were affected by its absence when using the SOI.

P2$_{ooss}$, who initially disliked how busy the overview was, reported how task strategies were affected by the loss of the overview; the main advantage of the overview for P2$_{ooss}$ was being able to hear the relative distance between stations. This is a sentiment that was echoed by P7$_{ssoo}$ who stated that the

overview *"doesn't say the exact location [of the station] but you actually can say how far relative the station is from the other station"*. P11$_{ssoo}$ used the overview not only for proximity but also to understand direction of travel: *"I have which direction to travel so the [transfer] station, I will choose the one [...] nearest to the station."*.

On the other hand, P16$_{ooss}$ specifically stated that the overview did not help with the proximity of stations. Instead, P16$_{ooss}$ stated that *"you can rely on [the overview]. You know it will pick up features [...] so you select the right way"*. P16$_{ooss}$ went on to describe the strategy for the SOI tasks saying *"I just picked a station and saw if it worked. The first task it did, the second task it didn't. And it got frustrating. [...] Without overviews you're totally lost"*.

This feedback confirms that the main overview and the information encoded in it proved advantageous in solving the tasks for these participants. Some mentioned being able to hear the relative location of stations, others optimising journeys by direction of travel. In other words, the overviews were used in a variety of ways. While not enough comments were made to report a clear preference, it is encouraging to hear how some put the overview to work for them.

B.3   Conclusion

Participant feedback on preferences for one interface over the other was limited. Those who stated as preference for the SOI talked about how the overviews were hard to understand; those who mentioned preferring or missing the overviews in the OAI referred to the added information that the overviews provided. So for the first camp, the overviews were a distraction and for the second, their absence increased their frustration. This implies that a participants understanding of the the overview drove whether they preferred the OAI over the SOI. It also seems that the amount of time to become comfortable with the non-speech audio interface differs per participant. For example, P2$_{ooss}$ changed opinion after experiencing the other interface and was more receptive after completing four tasks. It would be difficult to predict how long it would take a participant to get used to a novel interface however a potential solution would be to break up such evaluations into stages so that the participant has time to absorb the concepts at a more measured pace.

Appendix C

A log of a *meeting up* task

Below is a log from the auditory transportation system experiment described in Chapter 4. The logs were all encoded in XML as shown in Table 5.1. This log is from Participant 20 and is his solution to Task 3 using the overview-augmented interface. The first portion of the log describes the participant and the second describes the journeys the participant used in his solution. The remainder is the record of the steps the participant took to complete the task.

```
<log exp="20" cond="overview" task="3">
  <participant age="28" sex="M" compExp="15" hearingProbs="no"
      otherExperiments="RE diagrammes, motion capture">
    <metros>
      <metro name="london"/>
      <metro name="frankfurt"/>
      <metro name="budapest"/>
    </metros>
  </participant>
  <answer>
    <starts>
      <station dist="some">Harry</station>
      <station dist="some">X-Ray</station>
      <station dist="some|few">Zara</station>
    </starts>
    <meeting>Oscar</meeting>
    <transfer who="Zara">Baker</transfer>
    <other>Charlie</other>
  </answer>
  <nav time="0.002" key="start position" location="0 ">
      -- the menu
  </nav>
```

```
<nav time="6.069" key="down" location="0 0 ">
    -- the menu -- the legend
</nav>
<nav time="8.279" key="right" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="10.597" key="down" location="0 1 0 ">
    -- the menu -- the index -- Adam station
</nav>
<nav time="13.347" key="right" location="0 1 1 ">
    -- the menu -- the index -- Alpha station
</nav>
<nav time="17.875" key="right" location="0 1 2 ">
    -- the menu -- the index -- Ancona station
</nav>
<nav time="20.324" key="right" location="0 1 3 ">
    -- the menu -- the index -- Anton station
</nav>
<nav time="23.243" key="right" location="0 1 4 ">
    -- the menu -- the index -- Apple station
</nav>
<nav time="24.573" key="right" location="0 1 5 ">
    -- the menu -- the index -- Baker station
</nav>
<nav time="26.083" key="right" location="0 1 6 ">
    -- the menu -- the index -- Berta station
</nav>
<nav time="27.002" key="right" location="0 1 7 ">
    -- the menu -- the index -- Bravo station
</nav>
<nav time="28.441" key="right" location="0 1 8 ">
    -- the menu -- the index -- Caesar station
</nav>
<nav time="29.591" key="right" location="0 1 9 ">
    -- the menu -- the index -- Charlie station
</nav>
<nav time="31.07" key="right" location="0 1 10 ">
    -- the menu -- the index -- Charlotte station
</nav>
<nav time="32.62" key="right" location="0 1 11 ">
    -- the menu -- the index -- David station
</nav>
<nav time="34.21" key="right" location="0 1 12 ">
    -- the menu -- the index -- Delta station
</nav>
<nav time="35.73" key="right" location="0 1 13 ">
    -- the menu -- the index -- Desire station
</nav>
<nav time="37.25" key="right" location="0 1 14 ">
```

```
    -- the menu -- the index -- Dora station
</nav>
<nav time="38.868" key="right" location="0 1 15 ">
    -- the menu -- the index -- Echo station
</nav>
<nav time="40.328" key="right" location="0 1 16 ">
    -- the menu -- the index -- Edward station
</nav>
<nav time="41.838" key="right" location="0 1 17 ">
    -- the menu -- the index -- Fiji station
</nav>
<nav time="43.358" key="right" location="0 1 18 ">
    -- the menu -- the index -- Foxtrot station
</nav>
<nav time="44.966" key="right" location="0 1 19 ">
    -- the menu -- the index -- Freddy station
</nav>
<nav time="46.527" key="right" location="0 1 20 ">
    -- the menu -- the index -- Gustav station
</nav>
<nav time="48.115" key="right" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="52.484" key="left" location="0 1 20 ">
    -- the menu -- the index -- Gustav station
</nav>
<nav time="55.253" key="right" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="58.222" key="left" location="0 1 20 ">
    -- the menu -- the index -- Gustav station
</nav>
<nav time="59.182" key="right" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="63.05" key="down" location="0 1 21 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station
</nav>
<nav time="65.389" key="up" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="68.438" key="down" location="0 1 21 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station
</nav>
<nav time="70.167" key="up" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
```

```
<nav time="73.219" key="down" location="0 1 21 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station
</nav>
<nav time="80.754" key="down" location="0 1 21 0 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station -- red
</nav>
<nav time="81.454" key="up" location="0 1 21 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station
</nav>
<nav time="82.614" key="right" location="0 1 21 1 ">
    -- the menu -- the index -- Harry station
    -- filter on Harry station
</nav>
<filter time="85.103" type="station" which="Harry" state="on">
    -Harry-
</filter>
<nav time="99.188" key="up" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="100.048" key="left" location="0 1 20 ">
    -- the menu -- the index -- Gustav station
</nav>
<nav time="101.598" key="right" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="102.697" key="down" location="0 1 21 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station
</nav>
<nav time="104.307" key="down" location="0 1 21 0 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station -- red
</nav>
<nav time="107.217" key="right" location="0 1 21 0 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station -- red
</nav>
<nav time="108.646" key="up" location="0 1 21 0 ">
    -- the menu -- the index -- Harry station
    -- lines through Harry station
</nav>
<nav time="110.348" key="right" location="0 1 21 1 ">
    -- the menu -- the index -- Harry station
    -- turn off the filter on Harry station
</nav>
<nav time="111.414" key="right" location="0 1 21 2 ">
```

```
                -- the menu -- the index -- Harry station -- clear all filters
</nav>
<filter time="113.631" type="station" which="Harry" state="off"/>
<nav time="115.444" key="up" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="117.433" key="up" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="119.643" key="down" location="0 1 0 ">
    -- the menu -- the index -- Adam station
</nav>
<nav time="122.532" key="left" location="0 1 72 ">
    -- the menu -- the index -- Zulu station
</nav>
<nav time="123.842" key="left" location="0 1 71 ">
    -- the menu -- the index -- Zeppelin station
</nav>
<nav time="125.34" key="left" location="0 1 70 ">
    -- the menu -- the index -- Zebra station
</nav>
<nav time="126.99" key="left" location="0 1 69 ">
    -- the menu -- the index -- Zoe station
</nav>
<nav time="128.3" key="left" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="130.149" key="down" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="132.938" key="right" location="0 1 68 1 ">
    -- the menu -- the index -- Zara station
    -- filter on Zara station
</nav>
<filter time="136.369" type="station" which="Zara" state="on">
    -Zara-
</filter>
<nav time="144.668" key="up" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="145.524" key="left" location="0 1 67 ">
    -- the menu -- the index -- Yvonne station
</nav>
<nav time="146.763" key="right" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="147.853" key="down" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
```

```
        -- lines through Zara station
</nav>
<nav time="148.952" key="down" location="0 1 68 0 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station -- purple
</nav>
<nav time="150.523" key="right" location="0 1 68 0 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station -- purple
</nav>
<nav time="155.921" key="right" location="0 1 68 0 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station -- purple
</nav>
<nav time="157.673" key="down" location="0 1 68 0 0 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station -- purple
    -- stations on the purple line
</nav>
<nav time="157.92" key="up" location="0 1 68 0 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station -- purple
</nav>
<nav time="159.163" key="right" location="0 1 68 0 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station -- purple
</nav>
<nav time="161.469" key="up" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="163.099" key="right" location="0 1 68 1 ">
    -- the menu -- the index -- Zara station
    -- turn off the filter on Zara station
</nav>
<nav time="165.088" key="right" location="0 1 68 2 ">
    -- the menu -- the index -- Zara station
    -- clear all filters
</nav>
<filter time="166.158" type="station" which="Zara" state="off"/>
<nav time="167.868" key="up" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="169.037" key="left" location="0 1 67 ">
    -- the menu -- the index -- Yvonne station
</nav>
<nav time="169.797" key="up" location="0 1 ">
    -- the menu -- the index
</nav>
```

```
<nav time="171.867" key="down" location="0 1 0 ">
    -- the menu -- the index -- Adam station
</nav>
<nav time="174.225" key="left" location="0 1 72 ">
    -- the menu -- the index -- Zulu station
</nav>
<nav time="175.114" key="left" location="0 1 71 ">
    -- the menu -- the index -- Zeppelin station
</nav>
<nav time="175.195" key="left" location="0 1 72 ">
    -- the menu -- the index -- Zulu station
</nav>
<nav time="176.006" key="left" location="0 1 71 ">
    -- the menu -- the index -- Zeppelin station
</nav>
<nav time="176.904" key="left" location="0 1 70 ">
    -- the menu -- the index -- Zebra station
</nav>
<nav time="177.933" key="left" location="0 1 69 ">
    -- the menu -- the index -- Zoe station
</nav>
<nav time="178.884" key="left" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="179.923" key="left" location="0 1 67 ">
    -- the menu -- the index -- Yvonne station
</nav>
<nav time="181.217" key="left" location="0 1 66 ">
    -- the menu -- the index -- Yankee station
</nav>
<nav time="182.264" key="left" location="0 1 65 ">
    -- the menu -- the index -- X-Ray station
</nav>
<nav time="184.012" key="down" location="0 1 65 0 ">
    -- the menu -- the index -- X-Ray station
    -- lines through X-Ray station
</nav>
<nav time="185.641" key="down" location="0 1 65 0 0 ">
    -- the menu -- the index -- X-Ray station
    -- lines through X-Ray station -- blue
</nav>
<nav time="187.201" key="up" location="0 1 65 0 ">
    -- the menu -- the index -- X-Ray station
    -- lines through X-Ray station
</nav>
<nav time="198.66" key="right" location="0 1 65 1 ">
    -- the menu -- the index -- X-Ray station
    -- filter on X-Ray station
</nav>
```

```
<nav time="199.587" key="right" location="0 1 65 2 ">
    -- the menu -- the index -- X-Ray station
    -- clear all filters
</nav>
<nav time="201.047" key="up" location="0 1 65 ">
    -- the menu -- the index -- X-Ray station
</nav>
<nav time="202.146" key="right" location="0 1 66 ">
    -- the menu -- the index -- Yankee station
</nav>
<nav time="202.976" key="right" location="0 1 67 ">
    -- the menu -- the index -- Yvonne station
</nav>
<nav time="203.816" key="right" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="205.256" key="down" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="206.365" key="down" location="0 1 68 0 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station -- purple
</nav>
<nav time="212.664" key="up" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="213.934" key="up" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="215.653" key="up" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="217.032" key="left" location="0 0 ">
    -- the menu -- the legend
</nav>
<nav time="217.882" key="down" location="0 0 0 ">
    -- the menu -- the legend -- blue
</nav>
<nav time="220.78" key="right" location="0 0 1 ">
    -- the menu -- the legend -- green
</nav>
<nav time="222.251" key="right" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="225.46" key="down" location="0 0 2 0 ">
    -- the menu -- the legend -- purple
    -- stations on the purple line
```

```
</nav>
<nav time="231.397" key="right" location="0 0 2 1 ">
    -- the menu -- the legend -- purple
    -- filter on the purple line
</nav>
<filter time="233.754" type="line" which="purple" state="on">
    -purple-
</filter>
<nav time="243.703" key="up" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="244.894" key="up" location="0 0 ">
    -- the menu -- the legend
</nav>
<nav time="246.176" key="up" location="0 ">
    -- the menu
</nav>
<nav time="248.314" key="down" location="0 0 ">
    -- the menu -- the legend
</nav>
<nav time="249.643" key="right" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="250.751" key="down" location="0 1 0 ">
    -- the menu -- the index -- Adam station
</nav>
<nav time="253.8" key="left" location="0 1 72 ">
    -- the menu -- the index -- Zulu station
</nav>
<nav time="255.189" key="left" location="0 1 71 ">
    -- the menu -- the index -- Zeppelin station
</nav>
<nav time="256.339" key="left" location="0 1 70 ">
    -- the menu -- the index -- Zebra station
</nav>
<nav time="257.528" key="left" location="0 1 69 ">
    -- the menu -- the index -- Zoe station
</nav>
<nav time="258.638" key="left" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="260.148" key="down" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="262.368" key="right" location="0 1 68 1 ">
    -- the menu -- the index -- Zara station
    -- filter on Zara station
</nav>
```

```
<filter time="264.146" type="station" which="Zara" state="on">
    -Zara--purple-
</filter>
<nav time="265.807" key="up" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="266.886" key="up" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="269.038" key="left" location="0 0 ">
    -- the menu -- the legend
</nav>
<nav time="270.576" key="down" location="0 0 0 ">
    -- the menu -- the legend -- blue
</nav>
<nav time="271.795" key="down" location="0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
</nav>
<nav time="277.014" key="right" location="0 0 0 1 ">
    -- the menu -- the legend -- blue -- filter on the blue line
</nav>
<nav time="279.323" key="up" location="0 0 0 ">
    -- the menu -- the legend -- blue
</nav>
<nav time="279.792" key="right" location="0 0 1 ">
    -- the menu -- the legend -- green
</nav>
<nav time="280.322" key="right" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="281.751" key="down" location="0 0 2 0 ">
    -- the menu -- the legend -- purple
    -- stations on the purple line
</nav>
<nav time="283.6" key="down" location="0 0 2 0 0 ">
    -- the menu -- the legend -- purple
    -- stations on the purple line -- Nicolas station
</nav>
<nav time="284.977" key="up" location="0 0 2 0 ">
    -- the menu -- the legend -- purple
    -- stations on the purple line
</nav>
<nav time="305.656" key="up" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="310.463" key="left" location="0 0 1 ">
    -- the menu -- the legend -- green
</nav>
<nav time="312.082" key="left" location="0 0 0 ">
```

```
    -- the menu -- the legend -- blue
</nav>
<nav time="313.842" key="down" location="0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
</nav>
<nav time="318.59" key="right" location="0 0 0 1 ">
    -- the menu -- the legend -- blue -- filter on the blue line
</nav>
<nav time="318.614" key="right" location="0 0 0 2 ">
    -- the menu -- the legend -- blue -- clear all filters
</nav>
<filter time="320.319" type="station" which="Zara" state="off">
    -purple-
</filter>
<filter time="320.379" type="line" which="purple" state="off"/>
<nav time="327.027" key="up" location="0 0 0 ">
    -- the menu -- the legend -- blue
</nav>
<nav time="327.784" key="down" location="0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
</nav>
<nav time="330.006" key="right" location="0 0 0 1 ">
    -- the menu -- the legend -- blue -- filter on the blue line
</nav>
<filter time="331.978" type="line" which="blue" state="on">
    -blue-
</filter>
<nav time="342.712" key="up" location="0 0 0 ">
    -- the menu -- the legend -- blue
</nav>
<nav time="343.381" key="down" location="0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
</nav>
<nav time="345.911" key="down" location="0 0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
    -- Papa station
</nav>
<nav time="347.191" key="up" location="0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
</nav>
<nav time="347.888" key="right" location="0 0 0 1 ">
    -- the menu -- the legend -- blue
    -- turn off the filter on the blue line
</nav>
<nav time="350.869" key="right" location="0 0 0 2 ">
    -- the menu -- the legend -- blue -- clear all filters
</nav>
<nav time="352.41" key="right" location="0 0 0 3 ">
    -- the menu -- the legend -- blue -- go back
```

```
</nav>
<nav time="356.098" key="right" location="0 0 0 4 ">
    -- the menu -- the legend -- blue -- go home
</nav>
<nav time="357.488" key="right" location="0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
</nav>
<nav time="362.036" key="down" location="0 0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
    -- Papa station
</nav>
<nav time="367.074" key="up" location="0 0 0 0 ">
    -- the menu -- the legend -- blue -- stations on the blue line
</nav>
<nav time="380.58" key="right" location="0 0 0 1 ">
    -- the menu -- the legend -- blue
    -- turn off the filter on the blue line
</nav>
<nav time="381.377" key="up" location="0 0 0 ">
    -- the menu -- the legend -- blue
</nav>
<nav time="383" key="right" location="0 0 1 ">
    -- the menu -- the legend -- green
</nav>
<nav time="384.339" key="right" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="385.96" key="down" location="0 0 2 0 ">
    -- the menu -- the legend -- purple
    -- stations on the purple line
</nav>
<nav time="387.667" key="right" location="0 0 2 1 ">
    -- the menu -- the legend -- purple
    -- filter on the purple line
</nav>
<filter time="389.809" type="line" which="purple" state="on">
    -blue--purple-
</filter>
<nav time="417.788" key="up" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="422.059" key="down" location="0 0 2 0 ">
    -- the menu -- the legend -- purple
    -- stations on the purple line
</nav>
<nav time="425.507" key="up" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="430.984" key="up" location="0 0 ">
```

```
    -- the menu -- the legend
</nav>
<nav time="433.386" key="right" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="434.983" key="down" location="0 1 0 ">
    -- the menu -- the index -- Adam station
</nav>
<nav time="437.623" key="left" location="0 1 72 ">
    -- the menu -- the index -- Zulu station
</nav>
<nav time="438.212" key="left" location="0 1 71 ">
    -- the menu -- the index -- Zeppelin station
</nav>
<nav time="439.291" key="left" location="0 1 70 ">
    -- the menu -- the index -- Zebra station
</nav>
<nav time="440.381" key="left" location="0 1 69 ">
    -- the menu -- the index -- Zoe station
</nav>
<nav time="441.462" key="left" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="443.101" key="down" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="452.468" key="up" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="453.367" key="down" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="455.538" key="up" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="457.007" key="right" location="0 1 69 ">
    -- the menu -- the index -- Zoe station
</nav>
<nav time="459.236" key="left" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="460.432" key="down" location="0 1 68 0 ">
    -- the menu -- the index -- Zara station
    -- lines through Zara station
</nav>
<nav time="461.934" key="right" location="0 1 68 1 ">
    -- the menu -- the index -- Zara station
```

```
        -- filter on Zara station
</nav>
<filter time="463.923" type="station" which="Zara" state="on">
    -Zara--blue--purple-
</filter>
<nav time="512.94" key="up" location="0 1 68 ">
    -- the menu -- the index -- Zara station
</nav>
<nav time="514.718" key="up" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="516.37" key="left" location="0 0 ">
    -- the menu -- the legend
</nav>
<nav time="518.557" key="down" location="0 0 0 ">
    -- the menu -- the legend -- blue
</nav>
<nav time="520.388" key="right" location="0 0 1 ">
    -- the menu -- the legend -- green
</nav>
<nav time="521.756" key="right" location="0 0 2 ">
    -- the menu -- the legend -- purple
</nav>
<nav time="523.046" key="right" location="0 0 3 ">
    -- the menu -- the legend -- red
</nav>
<nav time="524.545" key="down" location="0 0 3 0 ">
    -- the menu -- the legend -- red -- stations on the red line
</nav>
<nav time="529.113" key="right" location="0 0 3 1 ">
    -- the menu -- the legend -- red -- filter on the red line
</nav>
<filter time="531.236" type="line" which="red" state="on">
    -Zara--blue--purple--red-
</filter>
<nav time="563.363" key="up" location="0 0 3 ">
    -- the menu -- the legend -- red
</nav>
<nav time="564.402" key="up" location="0 0 ">
    -- the menu -- the legend
</nav>
<nav time="565.635" key="right" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="566.804" key="down" location="0 1 0 ">
    -- the menu -- the index -- Adam station
</nav>
<nav time="568.351" key="down" location="0 1 0 0 ">
    -- the menu -- the index -- Adam station
```

```
        -- lines through Adam station
</nav>
<nav time="569.892" key="up" location="0 1 0 ">
        -- the menu -- the index -- Adam station
</nav>
<nav time="571.37" key="right" location="0 1 1 ">
        -- the menu -- the index -- Alpha station
</nav>
<nav time="572.342" key="right" location="0 1 2 ">
        -- the menu -- the index -- Ancona station
</nav>
<nav time="572.366" key="right" location="0 1 3 ">
        -- the menu -- the index -- Anton station
</nav>
<nav time="573.27" key="right" location="0 1 4 ">
        -- the menu -- the index -- Apple station
</nav>
<nav time="574.18" key="right" location="0 1 5 ">
        -- the menu -- the index -- Baker station
</nav>
<nav time="575.169" key="right" location="0 1 6 ">
        -- the menu -- the index -- Berta station
</nav>
<nav time="576.209" key="right" location="0 1 7 ">
        -- the menu -- the index -- Bravo station
</nav>
<nav time="577.269" key="right" location="0 1 8 ">
        -- the menu -- the index -- Caesar station
</nav>
<nav time="577.292" key="right" location="0 1 9 ">
        -- the menu -- the index -- Charlie station
</nav>
<nav time="578.268" key="right" location="0 1 10 ">
        -- the menu -- the index -- Charlotte station
</nav>
<nav time="580.668" key="right" location="0 1 11 ">
        -- the menu -- the index -- David station
</nav>
<nav time="581.707" key="left" location="0 1 10 ">
        -- the menu -- the index -- Charlotte station
</nav>
<nav time="582.856" key="down" location="0 1 10 0 ">
        -- the menu -- the index -- Charlotte station
        -- lines through Charlotte station
</nav>
<nav time="584.905" key="down" location="0 1 10 0 0 ">
        -- the menu -- the index -- Charlotte station
        -- lines through Charlotte station -- green
</nav>
```

```
<nav time="587.705" key="up" location="0 1 10 0 ">
    -- the menu -- the index -- Charlotte station
    -- lines through Charlotte station
</nav>
<nav time="588.425" key="right" location="0 1 10 1 ">
    -- the menu -- the index -- Charlotte station
    -- filter on Charlotte station
</nav>
<nav time="590.605" key="up" location="0 1 10 ">
    -- the menu -- the index -- Charlotte station
</nav>
<nav time="592.434" key="right" location="0 1 11 ">
    -- the menu -- the index -- David station
</nav>
<nav time="593.635" key="left" location="0 1 10 ">
    -- the menu -- the index -- Charlotte station
</nav>
<nav time="595.153" key="left" location="0 1 9 ">
    -- the menu -- the index -- Charlie station
</nav>
<nav time="596.874" key="down" location="0 1 9 0 ">
    -- the menu -- the index -- Charlie station
    -- lines through Charlie station
</nav>
<nav time="599.382" key="down" location="0 1 9 0 0 ">
    -- the menu -- the index -- Charlie station
    -- lines through Charlie station -- blue
</nav>
<nav time="601.482" key="right" location="0 1 9 0 1 ">
    -- the menu -- the index -- Charlie station
    -- lines through Charlie station -- purple
</nav>
<nav time="602.88" key="right" location="0 1 9 0 0 ">
    -- the menu -- the index -- Charlie station
    -- lines through Charlie station -- blue
</nav>
<nav time="604.36" key="right" location="0 1 9 0 1 ">
    -- the menu -- the index -- Charlie station
    -- lines through Charlie station -- purple
</nav>
<nav time="608.909" key="up" location="0 1 9 0 ">
    -- the menu -- the index -- Charlie station
    -- lines through Charlie station
</nav>
<nav time="624.404" key="up" location="0 1 9 ">
    -- the menu -- the index -- Charlie station
</nav>
<nav time="626.874" key="down" location="0 1 9 0 ">
    -- the menu -- the index -- Charlie station
```

```
        -- lines through Charlie station
</nav>
<nav time="629.762" key="up" location="0 1 9 ">
    -- the menu -- the index -- Charlie station
</nav>
<nav time="651.445" key="up" location="0 1 ">
    -- the menu -- the index
</nav>
<nav time="653.215" key="down" location="0 1 0 ">
    -- the menu -- the index -- Adam station
</nav>
<nav time="654.954" key="right" location="0 1 1 ">
    -- the menu -- the index -- Alpha station
</nav>
<nav time="655.823" key="right" location="0 1 2 ">
    -- the menu -- the index -- Ancona station
</nav>
<nav time="656.715" key="right" location="0 1 3 ">
    -- the menu -- the index -- Anton station
</nav>
<nav time="659.164" key="right" location="0 1 4 ">
    -- the menu -- the index -- Apple station
</nav>
<nav time="660.154" key="right" location="0 1 5 ">
    -- the menu -- the index -- Baker station
</nav>
<nav time="661.132" key="right" location="0 1 6 ">
    -- the menu -- the index -- Berta station
</nav>
<nav time="662.081" key="right" location="0 1 7 ">
    -- the menu -- the index -- Bravo station
</nav>
<nav time="663.072" key="right" location="0 1 8 ">
    -- the menu -- the index -- Caesar station
</nav>
<nav time="664.021" key="right" location="0 1 9 ">
    -- the menu -- the index -- Charlie station
</nav>
<nav time="665.02" key="right" location="0 1 10 ">
    -- the menu -- the index -- Charlotte station
</nav>
<nav time="665.96" key="right" location="0 1 11 ">
    -- the menu -- the index -- David station
</nav>
<nav time="666.9" key="right" location="0 1 12 ">
    -- the menu -- the index -- Delta station
</nav>
<nav time="667.82" key="right" location="0 1 13 ">
    -- the menu -- the index -- Desire station
```

```
</nav>
<nav time="668.85" key="right" location="0 1 14 ">
    -- the menu -- the index -- Dora station
</nav>
<nav time="669.789" key="right" location="0 1 15 ">
    -- the menu -- the index -- Echo station
</nav>
<nav time="670.649" key="right" location="0 1 16 ">
    -- the menu -- the index -- Edward station
</nav>
<nav time="671.549" key="right" location="0 1 17 ">
    -- the menu -- the index -- Fiji station
</nav>
<nav time="672.558" key="right" location="0 1 18 ">
    -- the menu -- the index -- Foxtrot station
</nav>
<nav time="673.638" key="right" location="0 1 19 ">
    -- the menu -- the index -- Freddy station
</nav>
<nav time="674.788" key="right" location="0 1 20 ">
    -- the menu -- the index -- Gustav station
</nav>
<nav time="677.477" key="right" location="0 1 21 ">
    -- the menu -- the index -- Harry station
</nav>
<nav time="678.617" key="right" location="0 1 22 ">
    -- the menu -- the index -- Hotel station
</nav>
<nav time="679.866" key="right" location="0 1 23 ">
    -- the menu -- the index -- Ida station
</nav>
<nav time="680.908" key="right" location="0 1 24 ">
    -- the menu -- the index -- Item station
</nav>
<nav time="681.845" key="right" location="0 1 25 ">
    -- the menu -- the index -- India station
</nav>
<nav time="682.815" key="right" location="0 1 26 ">
    -- the menu -- the index -- John station
</nav>
<nav time="683.855" key="right" location="0 1 27 ">
    -- the menu -- the index -- Joseph station
</nav>
<nav time="684.855" key="right" location="0 1 28 ">
    -- the menu -- the index -- Juliet station
</nav>
<nav time="685.835" key="right" location="0 1 29 ">
    -- the menu -- the index -- Kilo station
</nav>
```

```
<nav time="687.104" key="right" location="0 1 30 ">
    -- the menu -- the index -- Leema station
</nav>
<nav time="688.284" key="right" location="0 1 31 ">
    -- the menu -- the index -- Louis station
</nav>
<nav time="689.563" key="right" location="0 1 32 ">
    -- the menu -- the index -- Marcel station
</nav>
<nav time="691.153" key="right" location="0 1 33 ">
    -- the menu -- the index -- Martha station
</nav>
<nav time="692.433" key="right" location="0 1 34 ">
    -- the menu -- the index -- Mary station
</nav>
<nav time="693.392" key="right" location="0 1 35 ">
    -- the menu -- the index -- Mother station
</nav>
<nav time="694.312" key="right" location="0 1 36 ">
    -- the menu -- the index -- Nicolas station
</nav>
<nav time="695.282" key="right" location="0 1 37 ">
    -- the menu -- the index -- November station
</nav>
<nav time="696.252" key="right" location="0 1 38 ">
    -- the menu -- the index -- Oboe station
</nav>
<nav time="697.191" key="right" location="0 1 39 ">
    -- the menu -- the index -- Ocean station
</nav>
<nav time="698.69" key="right" location="0 1 40 ">
    -- the menu -- the index -- Orange station
</nav>
<nav time="700.285" key="right" location="0 1 41 ">
    -- the menu -- the index -- Oscar station
</nav>
<nav time="701.97" key="down" location="0 1 41 0 ">
    -- the menu -- the index -- Oscar station
    -- lines through Oscar station
</nav>
<nav time="703.869" key="down" location="0 1 41 0 0 ">
    -- the menu -- the index -- Oscar station
    -- lines through Oscar station -- blue
</nav>
<nav time="708.417" key="up" location="0 1 41 0 ">
    -- the menu -- the index -- Oscar station
    -- lines through Oscar station
</nav>
<nav time="709.176" key="down" location="0 1 41 0 0 ">
```

```
            -- the menu -- the index -- Oscar station
            -- lines through Oscar station -- blue
</nav>
<nav time="710.336" key="right" location="0 1 41 0 1 ">
            -- the menu -- the index -- Oscar station
            -- lines through Oscar station -- red
</nav>
<nav time="711.566" key="right" location="0 1 41 0 0 ">
            -- the menu -- the index -- Oscar station
            -- lines through Oscar station -- blue
</nav>
<nav time="741.978" key="down" location="0 1 41 0 0 0 ">
            -- the menu -- the index -- Oscar station
            -- lines through Oscar station -- blue
            -- stations on the blue line
</nav>
<nav time="744.676" key="up" location="0 1 41 0 0 ">
            -- the menu -- the index -- Oscar station
            -- lines through Oscar station -- blue
</nav>
<nav time="746.536" key="up" location="0 1 41 0 ">
            -- the menu -- the index -- Oscar station
            -- lines through Oscar station
</nav>
<nav time="748.646" key="right" location="0 1 41 1 ">
            -- the menu -- the index -- Oscar station
            -- filter on Oscar station
</nav>
<filter time="750.534" type="station" which="Oscar" state="on">
            -Zara--Oscar--blue--purple--red-
</filter>
<nav time="761" key="up" location="0 1 41 ">
            -- the menu -- the index -- Oscar station
</nav>
<nav time="834.947" key="up" location="0 1 ">
            -- the menu -- the index
</nav>
<nav time="837.286" key="down" location="0 1 0 ">
            -- the menu -- the index -- Adam station
</nav>
<nav time="838.447" key="right" location="0 1 1 ">
            -- the menu -- the index -- Alpha station
</nav>
<nav time="838.817" key="right" location="0 1 2 ">
            -- the menu -- the index -- Ancona station
</nav>
<nav time="839.206" key="right" location="0 1 3 ">
            -- the menu -- the index -- Anton station
</nav>
```

```
<nav time="839.766" key="right" location="0 1 4 ">
    -- the menu -- the index -- Apple station
</nav>
<nav time="840.575" key="right" location="0 1 5 ">
    -- the menu -- the index -- Baker station
</nav>
<nav time="841.495" key="right" location="0 1 6 ">
    -- the menu -- the index -- Berta station
</nav>
<nav time="842.375" key="right" location="0 1 7 ">
    -- the menu -- the index -- Bravo station
</nav>
<nav time="843.255" key="right" location="0 1 8 ">
    -- the menu -- the index -- Caesar station
</nav>
<nav time="844.175" key="right" location="0 1 9 ">
    -- the menu -- the index -- Charlie station
</nav>
<nav time="845.055" key="right" location="0 1 10 ">
    -- the menu -- the index -- Charlotte station
</nav>
<nav time="845.905" key="right" location="0 1 11 ">
    -- the menu -- the index -- David station
</nav>
<nav time="846.735" key="right" location="0 1 12 ">
    -- the menu -- the index -- Delta station
</nav>
<nav time="847.585" key="right" location="0 1 13 ">
    -- the menu -- the index -- Desire station
</nav>
<nav time="848.453" key="right" location="0 1 14 ">
    -- the menu -- the index -- Dora station
</nav>
<nav time="849.283" key="right" location="0 1 15 ">
    -- the menu -- the index -- Echo station
</nav>
<nav time="850.072" key="right" location="0 1 16 ">
    -- the menu -- the index -- Edward station
</nav>
<nav time="850.883" key="right" location="0 1 17 ">
    -- the menu -- the index -- Fiji station
</nav>
<nav time="851.823" key="right" location="0 1 18 ">
    -- the menu -- the index -- Foxtrot station
</nav>
<nav time="852.781" key="right" location="0 1 19 ">
    -- the menu -- the index -- Freddy station
</nav>
<nav time="853.752" key="right" location="0 1 20 ">
```

```
        -- the menu -- the index -- Gustav station
</nav>
<nav time="854.711" key="right" location="0 1 21 ">
        -- the menu -- the index -- Harry station
</nav>
<nav time="855.621" key="right" location="0 1 22 ">
        -- the menu -- the index -- Hotel station
</nav>
<nav time="856.531" key="left" location="0 1 21 ">
        -- the menu -- the index -- Harry station
</nav>
<nav time="857.69" key="down" location="0 1 21 0 ">
        -- the menu -- the index -- Harry station
        -- lines through Harry station
</nav>
<nav time="859.572" key="right" location="0 1 21 1 ">
        -- the menu -- the index -- Harry station
        -- filter on Harry station
</nav>
<nav time="861.309" key="left" location="0 1 21 0 ">
        -- the menu -- the index -- Harry station
        -- lines through Harry station
</nav>
<nav time="862.41" key="down" location="0 1 21 0 0 ">
        -- the menu -- the index -- Harry station
        -- lines through Harry station -- red
</nav>
<nav time="863.479" key="right" location="0 1 21 0 0 ">
        -- the menu -- the index -- Harry station
        -- lines through Harry station -- red
</nav>
<nav time="864.598" key="up" location="0 1 21 0 ">
        -- the menu -- the index -- Harry station
        -- lines through Harry station
</nav>
<nav time="867.967" key="right" location="0 1 21 1 ">
        -- the menu -- the index -- Harry station
        -- filter on Harry station
</nav>
<filter time="869.966" type="station" which="Harry" state="on">
        -Zara--Oscar--Harry--blue--purple--red-
</filter>
<nav time="977.543" key="up" location="0 1 21 ">
        -- the menu -- the index -- Harry station
</nav>
<nav time="979.572" key="right" location="0 1 22 ">
        -- the menu -- the index -- Hotel station
</nav>
<nav time="980.871" key="right" location="0 1 23 ">
```

```
        -- the menu -- the index -- Ida station
</nav>
<nav time="981.781" key="right" location="0 1 24 ">
    -- the menu -- the index -- Item station
</nav>
<nav time="982.681" key="right" location="0 1 25 ">
    -- the menu -- the index -- India station
</nav>
<nav time="983.581" key="right" location="0 1 26 ">
    -- the menu -- the index -- John station
</nav>
<nav time="984.501" key="right" location="0 1 27 ">
    -- the menu -- the index -- Joseph station
</nav>
<nav time="985.44" key="right" location="0 1 28 ">
    -- the menu -- the index -- Juliet station
</nav>
<nav time="986.31" key="right" location="0 1 29 ">
    -- the menu -- the index -- Kilo station
</nav>
<nav time="987.239" key="right" location="0 1 30 ">
    -- the menu -- the index -- Leema station
</nav>
<nav time="988.109" key="right" location="0 1 31 ">
    -- the menu -- the index -- Louis station
</nav>
<nav time="988.999" key="right" location="0 1 32 ">
    -- the menu -- the index -- Marcel station
</nav>
<nav time="989.869" key="right" location="0 1 33 ">
    -- the menu -- the index -- Martha station
</nav>
<nav time="990.768" key="right" location="0 1 34 ">
    -- the menu -- the index -- Mary station
</nav>
<nav time="991.708" key="right" location="0 1 35 ">
    -- the menu -- the index -- Mother station
</nav>
<nav time="992.587" key="right" location="0 1 36 ">
    -- the menu -- the index -- Nicolas station
</nav>
<nav time="993.418" key="right" location="0 1 37 ">
    -- the menu -- the index -- November station
</nav>
<nav time="994.277" key="right" location="0 1 38 ">
    -- the menu -- the index -- Oboe station
</nav>
<nav time="995.128" key="right" location="0 1 39 ">
    -- the menu -- the index -- Ocean station
```

```
</nav>
<nav time="995.977" key="right" location="0 1 40 ">
    -- the menu -- the index -- Orange station
</nav>
<nav time="996.898" key="right" location="0 1 41 ">
    -- the menu -- the index -- Oscar station
</nav>
<nav time="997.836" key="right" location="0 1 42 ">
    -- the menu -- the index -- Otto station
</nav>
<nav time="998.716" key="right" location="0 1 43 ">
    -- the menu -- the index -- Papa station
</nav>
<nav time="999.615" key="right" location="0 1 44 ">
    -- the menu -- the index -- Paula station
</nav>
<nav time="1000.536" key="right" location="0 1 45 ">
    -- the menu -- the index -- Peter station
</nav>
<nav time="1001.445" key="right" location="0 1 46 ">
    -- the menu -- the index -- Quebec station
</nav>
<nav time="1002.295" key="right" location="0 1 47 ">
    -- the menu -- the index -- Raoul station
</nav>
<nav time="1003.124" key="right" location="0 1 48 ">
    -- the menu -- the index -- Richard station
</nav>
<nav time="1003.994" key="right" location="0 1 49 ">
    -- the menu -- the index -- Robert station
</nav>
<nav time="1004.824" key="right" location="0 1 50 ">
    -- the menu -- the index -- Roger station
</nav>
<nav time="1005.674" key="right" location="0 1 51 ">
    -- the menu -- the index -- Romeo station
</nav>
<nav time="1006.584" key="right" location="0 1 52 ">
    -- the menu -- the index -- Sierra station
</nav>
<nav time="1007.463" key="right" location="0 1 53 ">
    -- the menu -- the index -- Sugar station
</nav>
<nav time="1009.133" key="right" location="0 1 54 ">
    -- the menu -- the index -- Suzanne station
</nav>
<nav time="1010.202" key="right" location="0 1 55 ">
    -- the menu -- the index -- Tango station
</nav>
```

```
<nav time="1011.442" key="right" location="0 1 56 ">
    -- the menu -- the index -- Theresa station
</nav>
<nav time="1012.362" key="right" location="0 1 57 ">
    -- the menu -- the index -- Uncle station
</nav>
<nav time="1013.271" key="right" location="0 1 58 ">
    -- the menu -- the index -- Uniform station
</nav>
<nav time="1014.241" key="right" location="0 1 59 ">
    -- the menu -- the index -- Union station
</nav>
<nav time="1015.141" key="right" location="0 1 60 ">
    -- the menu -- the index -- Victor station
</nav>
<nav time="1016.031" key="right" location="0 1 61 ">
    -- the menu -- the index -- Washington station
</nav>
<nav time="1017.01" key="right" location="0 1 62 ">
    -- the menu -- the index -- Whiskey station
</nav>
<nav time="1018" key="right" location="0 1 63 ">
    -- the menu -- the index -- William station
</nav>
<nav time="1019.011" key="right" location="0 1 64 ">
    -- the menu -- the index -- Xavier station
</nav>
<nav time="1019.969" key="right" location="0 1 65 ">
    -- the menu -- the index -- X-Ray station
</nav>
<nav time="1021.781" key="down" location="0 1 65 0 ">
    -- the menu -- the index -- X-Ray station
    -- lines through X-Ray station
</nav>
<nav time="1023.578" key="right" location="0 1 65 1 ">
    -- the menu -- the index -- X-Ray station
    -- filter on X-Ray station
</nav>
<filter time="1025.688" type="station" which="X-Ray" state="on">
    -Zara--Oscar--Harry--X-Ray--blue--purple--red-
</filter>
</log>
```

Appendix D

An analysis of *meeting up* tasks with a score of zero

There were five tasks with a score of zero from the auditory public transportation system application experiment (see Chapter 4). This appendix is an investigation into what the problems were.

D.1 Experiment 8, Task 2

The starting stations for Task 2 are Bravo (Purple line), Uniform (Blue line) and Washington (Green line). The participant used the overview-augmented interface.

This particular task took the participant just under 4.5 minutes. This was the fastest completion time from any participant. The following summarises the steps the participant took to solve the task.

Step 1 (0:00) participant turned off filters from the previous task.

Step 2 (0:18) participant navigated to Bravo station in the Index, investigated what line it was on and set a filter on the Purple line.

Step 3 (1:20) participant navigated to Washington station in the Index, investigated what line it was on and set a filter on the Green line.

Step 4 (2:31) participant navigate it to Uniform station in the Index, investigated what line it was on and set a filter on the Blue line.

Step 5 (3:11) participant listened to the main overview for 1:18.

When the participant applied the first filter (Step 2), the main overview would have played only the stations on the Purple line. When the participant applied the second filter (Step 3), the main overview

would have played the Purple line and the Green line and would have highlighted Delta station. With the addition of the third filter (Step 4), the main overview would have played the Purple line, the Green line and the Blue line; it would have also highlighted Delta station, Charlie station and Raoul station. The participant spent over a minute listening to the resulting overview.

With this little information, it is difficult to determine what the participant was thinking. The proposed solution to the task was that the meeting place be Bravo station. The person coming from Bravo station would stay in place, the person coming from Uniform station would travel directly to Bravo and the person coming from Washington station would travel to Bravo via Charlie station. Two of these journeys are impossible: Uniform cannot travel directly to Bravo as they are on different lines and Washington cannot use Charlie is a transfer station as they are also on different lines.

The speed with which the participant completed the task indicates that the participant rushed through the task and mistakes were made. One reasonable deduction is that the participant made two mistakes. The first mistake was that the participant confused Washington station with Uniform station. The second mistake was that the participant confused Bravo station and Delta station when writing down the answer. By fixing these two mistakes, the journeys become possible and reasonable: Bravo and Washington travel directly to Delta and Uniform travels to Delta via Charlie station, as per Figure 5.2. Figure D.1 shows the participant's proposed solution to Task 2 and the deduced intended solution.

D.2   Experiment 8, Task 4

The starting stations for Task 4 are Freddy (Red line), Marcel (Yellow line) and Uncle (blue line). The participant used the speech-only interface.

Here, the participant took approximately 30 minutes to complete the task. The solution is viable except for one journey where the participant used a simple station as the transfer point. Their solution has the person travelling from Freddy station going to Oscar to transfer to the blue line to meet at Charlie; the person at Uncle travels directly to Charlie; the person travelling from Marcel is directed to Uniform station (which is on a different line to and is not transfer station) and then on to Charlie. Following are the steps the participant took to come to their solution:

Step 1 (0:00) participant navigated to Freddy station in the Index and investigated what line it was on.

Step 2 (1:00) participant navigated to Marcel station in the Index and investigated what line it was on.

Step 3 (1:43) participant navigated to Uncle station in the Index and investigated what line it was on.
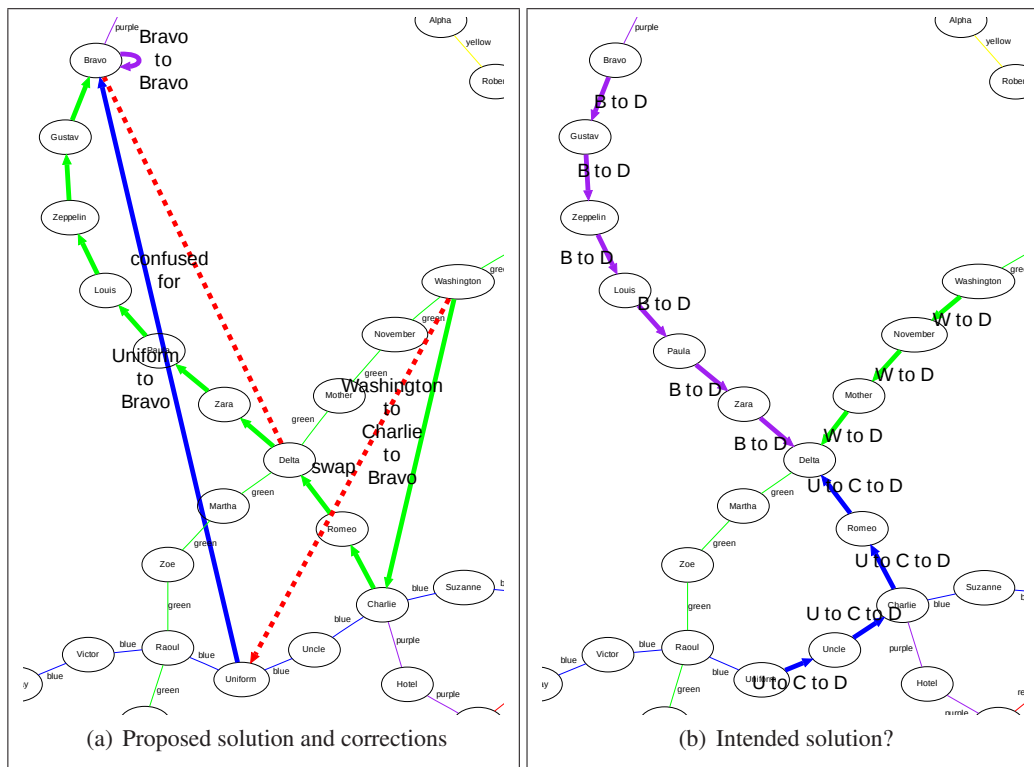
Figure D.1: **Correction of the solution from Experiment 8, Task 2.** D.1(a) shows the solution as proposed by the participant. The purple, blue and green lines show the journeys, two of which are impossible. The dashed red lines show the likely mistakes: the participant wrote Bravo into the solution instead of Delta and got Uniform and Washington reversed. D.1(b) shows the resulting journeys if those mistakes are corrected.

Step 4 (3:43) participant navigated to the Blue line in the Legend and listed its stations several times, including Uniform, Uncle, Charlie and Oscar.

Step 5 (6:22) participant navigated to the Red line in the Legend and listed its stations, including Freddy and Oscar.

Step 6 (7:38) participant navigated to the Yellow line in the Legend and listed its stations twice, including Marcel.

Step 7 (9:41) participant navigated to the Blue line in the Legend and listed its stations, including Uniform, Uncle, Charlie and Oscar.

Step 8 (10:27) participant navigated to the Red line in the Legend and listed its stations, including Freddy and Oscar.

Step 9 (11:10) participant navigated to the Yellow line in the Legend and listed its stations backwards, including Marcel.

Step 10 (15:12) participant navigated to the Red line in the Legend and listed its stations twice, including Freddy and Oscar.

Step 11 (16:27) participant navigated to Freddy in the Index and investigated what line it was on.

Step 12 (16:56) participant navigated to the Red line in the Legend, listed its first 7 stations, including Freddy and Oscar, and then paused for 12:35 to write down the answer.

step 13 (30:23) participant navigated to the Blue line in the Legend and listed its first eight stations, including Uniform, Uncle and Charlie.
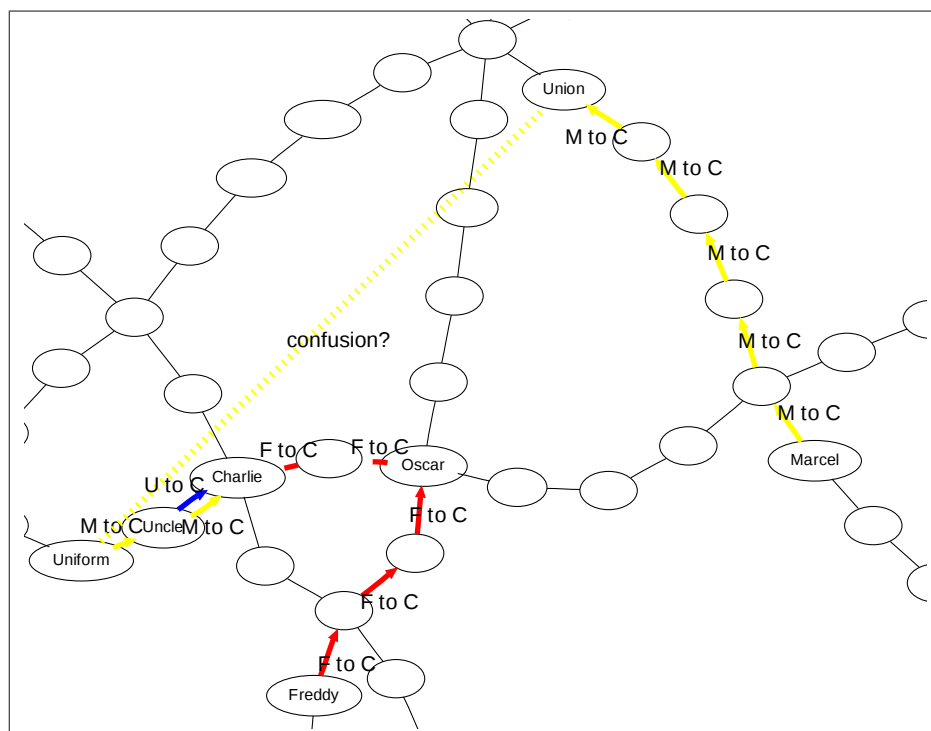
Figure D.2: **Correction of the solution from Experiment 8, Task 4.** The journeys specified by the participant from Freddy and Uncle are shown in red and blue respectively. The yellow shows the probable confusion: the participant mistook Union for Uniform and constructed an impossible journey for Marcel.

The choice of Charlie as the meeting point seems rather arbitrary unless one assumes that the participant confused Uniform station with Union station, which is on the Yellow line, same as Marcel station. From the steps described above, it seems that the participant formulated a solution during Step 12 and Step 13 is a re-confirmation of information. By then the stations on the Blue and Red lines were repeated several times (steps 4, 5, 7, 8, 10 and 12) and it is logical to conclude that that the participant understood that Oscar station was the transfer point between the two lines. This resulted in the journey from Freddy

station (Freddy → Oscar → Charlie). The participant seemed to be aware that Charlie and Uncle were on the same line (steps 3, 7 and 13), resulting in the journey from Uncle station (Uncle → Charlie). The only logical explanation for the journey from Marcel (Marcel → Uniform → Charlie) is that the participant confused Uniform station with Union. The reason I've come to this conclusion is that the participants solely relied on hearing the names of the stations spoken out: at no point did the participant explore what are lines went through the transfer stations that were used in the solution. As such, there is nothing that would have indicated to the participant that Union station and Uniform station were not the same. Figure D.2 shows the journeys specified by the participant.

D.3   Experiment 17, Task 1

The starting stations for Task 1 are Ida (Yellow line), November (Green line) and Zeppelin (Purple line). The participant used the speech-only interface.

The participant's solution to the task was to have the people coming from Ida and November stations travel directly to Lima and the person from Zeppelin station travel to Lima station using Charlie as a transfer point. The first two journeys, from Ida and November, are possible however, while one can travel from Zeppelin to Charlie, the journey between Charlie and Lima would require another transfer. This journey can be seen in Figure D.3. Following other steps the participant up to reach the solution:

Step 1 (0:00) participant navigated to Ida station in the Index and investigated what line it was on.

Step 2 (1: 42) participant navigated to Zeppelin station in the Index, investigated what line it was on and listed the stations on the line, including Fiji, Zeppelin and Charlie.

Step 3 (3:38) participant navigated to November station in the Index, investigated what line it was on and listed the stations on the line, including Lima and November.

Step 4 (6:21) participant navigated to the Green line in the Legend and listed all its stations including Lima and November.

Step 5 (9:56) participant navigated to Zeppelin station in the Index and listed various subsets of its stations (stations 1-11, 1-6, 6-1 and 16-10), including Fiji, Zeppelin and Charlie.

Step 6 (11:33) participant navigated to the Green line in the Legend, listed various stations on the line (including November) and searched for transfer stations on either side of November (no transfer stations found).
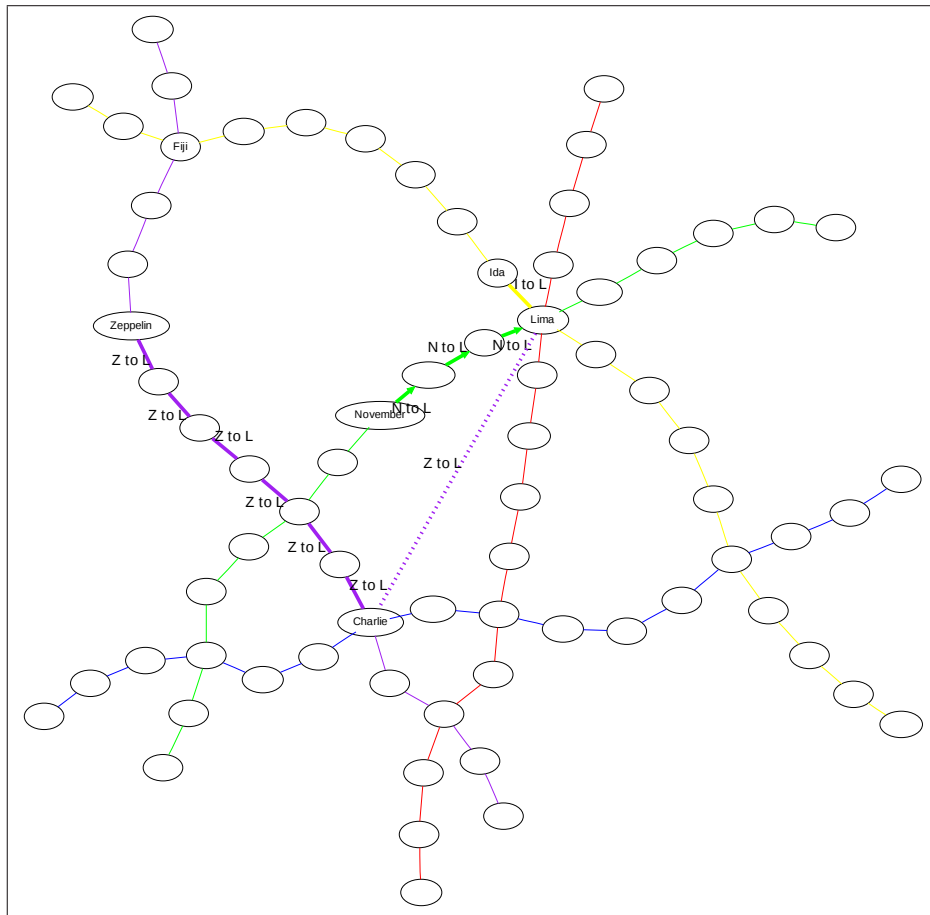
Figure D.3: **The solution from Experiment 17, Task 1.** The participant's solution to Task 1 sent the person at Ida station directly to Lima (Yellow line), the person at November station directly to Lima (Green line) and the person at Zeppelin station (Purple line) to Charlie and then on to Lima. It is this last leg that is impossible.

Step 7 (15:50) participant navigated to the Purple line in the Legend and searched the last four stations for transfer stations, finding no stations of interest.

Step 8 (17:43) participant navigated to the Blue line in the Legend and listed all the stations on the line, including Charlie.

Step 9 (20:45) participant navigated to the Purple line in the Legend and the listed all the stations on the line twice, including Fiji, Zeppelin and Charlie.

Step 10 (23:06) participant navigated to the Yellow line in the Legend and listed all the stations on the line twice, including Lima, Ida and Fiji.

Step 11 (25:43) participant navigated to the Green line in the Legend and listed all its stations twice, including Lima and November.

Step 12 (29:47) participant navigated to the Yellow line in the Legend and listed the first 10 stations on the line, including Lima.

In this task, the participant focused first on figuring out where November and Zeppelin are. Steps 2 through 7 above are all explorations of the Green and Purple lines. It seems that the participant first wanted to work out how to transfer between these two lines. Then in Step 8, the participant explores the Blue line; it is not clear what the rationale was here. The last three steps are the participant factoring the Yellow line (Ida station) into the solution. There are three possibilities as to why the participant made a mistake in the solution: first, the participant could have confused Charlie and Fiji stations; second, when exploring the blue line, the participant could have thought they were exploring the Green line; third, the participant was confused and chose a solution at random.

Confusing Charlie for Fiji

The participant was solely relying on hearing the names of stations in order to determine if they were on several lines and it is possible that when the participant explored the Yellow line, Fiji sounded like Charlie. This would resolve the impossible journey nicely. Charlie and Fiji both end in the same sound. However the similarity ends there. The beginning sounds of both Charlie and Fiji are quite different and this possibility seems slightly far-fetched.

Confusing the Green line for the Blue line

At the beginning of the task, the participant established a loose pattern of alternately exploring the Green and Purple lines. This pattern breaks in Step 8 when the participant explores the Blue line. Up until that point, the pattern was: Purple-Green-Green-Purple-Green-Purple. The participant then explores the Blue line and hears Charlie station, which was recognised from the explorations of the Purple line. At that point, the participant moves on to figure out how the Yellow line fit in with the other two lines. Having determined that Ida and Lima lie on the same line, the participant double checks that both November and Lima can travel directly to Lima and submits their solution to the task. If this were the case, that the participant thought they were exploring the Green line when they were in fact exploring Blue line, it would explain why the participant sent the person from Zeppelin station via Charlie. However there are two things that dispute this theory: first, when exploring the blue line, the participant would not have heard November station which should have been expected; second, when navigating to the Blue line, the participant did so at a very leisurely pace (54s between *The Blue line* and

Stations on the Blue line and a further 10s until *Papa station*, the first station on the Blue line). Despite this, I believe this possibility to be the most likely.

General confusion

The third possibility is that the participant was in general confused. This was their first task and they were just getting first hand experience with the application. While the interaction indicates that are the participant had a general idea of how to go about the task, I have no way of knowing how confident they were in their solution. It could be that they had not made sense of the task yet, that they were confused about the organisation of the application or that they had trouble understanding the text-to-speech other than in the ways I described above.

D.4 Experiment 18, Task 2

The starting stations for Task 2 are Bravo (Purple line), Uniform (Blue line) and Washington (Green line). The participant used the overview-augmented interface.

The participant's solution was that the people coming from Bravo and Washington stations travel directly to Fiji and the person coming from Uniform travel to Fiji via Charlie station. While the journeys for Bravo and Uniform stations are valid, Washington station is on the Green line and cannot travel directly to Fiji, which is on the Yellow and Purple lines.

Step 1 (0:00) participant navigated to Uniform station in the Index, investigated what line it was on (Blue) and set a filter on the Blue line.

Step 2 (02:41) participant navigated to Bravo station in the Index, set a filter on Bravo station, investigated what line it was on and set a filter on the Purple line.

Step 3 (06:15) participant navigated to Washington station in the Index and moved back and forth several times between Whiskey station and Washington station.

Step 4 (09:38) participant navigated to Whiskey station, set a filter on Whiskey station, investigated what line it was on (Yellow line) and paused to listen for 05:37.

Step 5 (18:48) participant navigated back to Whiskey station in the Index, investigated what line it was on and discovered that Whiskey is the first station on the Yellow line.
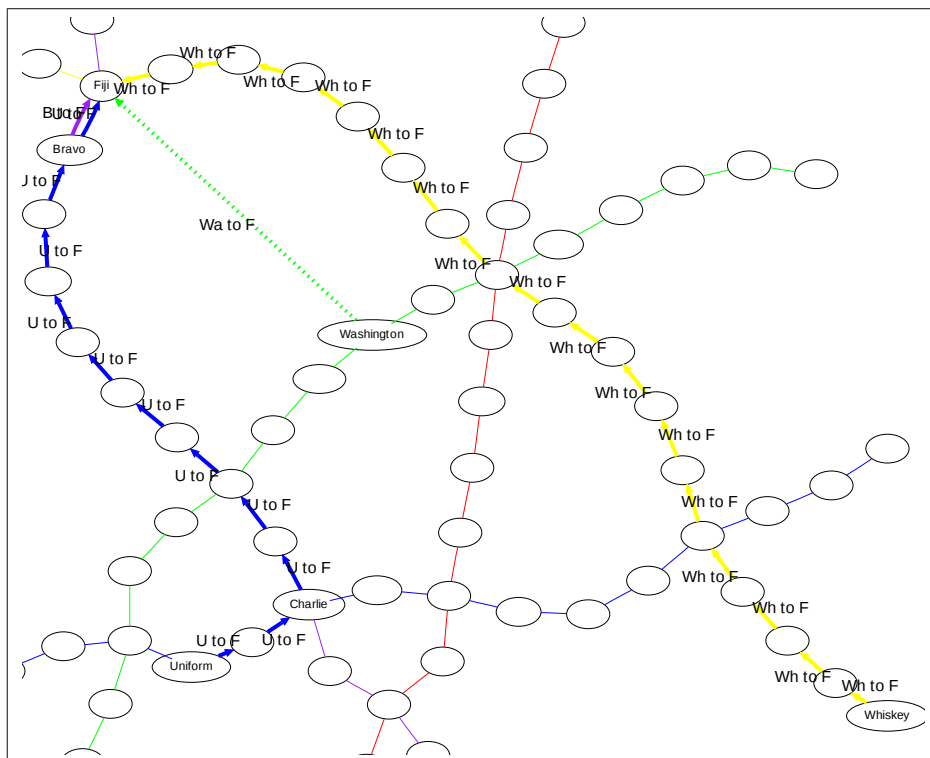
Figure D.4: **The solution from Experiment 18, Task 2.** The blue and purple lines in this diagram show the journeys for the people coming from Bravo station and Uniform station. The green dotted line shows a solution put forward by the participant however this journey is not possible. Steps 3 and 4 indicate that the participant confused on the text-to-speech for Whiskey for Washington. The yellow line shows the journey of a person coming from Whiskey station.

Step 6 (23:57) participant navigated to Charlie station in the Index and investigated what lines it is on (Blue and Purple).

Step 7 (25:06) participant navigated to Sierra station and investigated what line it was on (Blue).

It is fairly obvious that in Step 3, to the participant was having trouble with the pronunciation of Washington. By Step 4, the participant had decided that Whiskey was the correct station. Taking this into account, the participant solution is valid. One can travel directly from Whiskey to Fiji and the other two journeys remain the same. This is shown in Figure D.4.

D.5 Experiment 19, Task 3

The starting stations for Task 3 are Harry (Red line), X-Ray (Blue line) and Zara (Purple line). The participant used the overview-augmented interface.

The participant's solution had the people at Harry and X-Ray stations travel directly to Oscar station. The third person, coming from Zara station, was to also travel directly to Oscar. Zara station which is

on the Purple line cannot or do that without transferring. The following shows the steps the participant took:
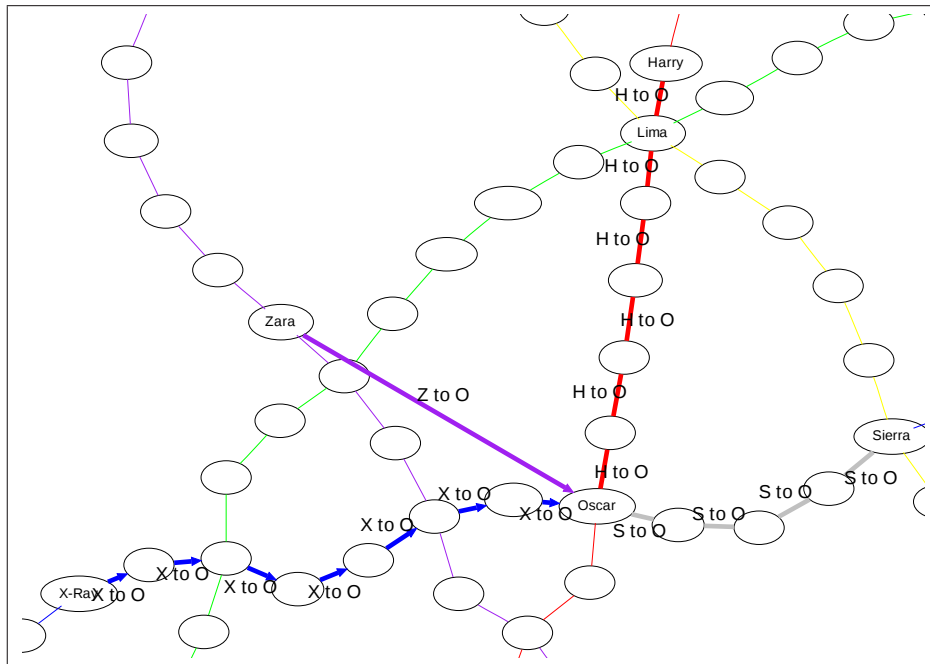
Figure D.5: **The solution from Experiment 19, Task 3.** The red and blue lines show the journeys from Harry and X-Ray to Oscar. The purple line shows the impossible journey from Zara to Oscar. Had the participant confused Sierra for Zara, the grey line shows the journey the participant intended.

Step 1 (0:00) participant navigated to the Blue line in the Legend, started listing stations on the Blue line, encountered X-Ray, set a filter on X-Ray, listed various stations on the Blue line, searched various stations for transfers and set a filter on the Blue line.

Step 2 (05:49) participant navigated to the Green line in the Legend and listed all its stations.

Step 3 (07:54) participant navigated to the Purple line and the Legend and listed all its stations, pausing around Zara station.

Step 4 (08:49) participant navigated to the Red line in the Legend, listed all stations up to Harry, explored Harry station, set a filter on the Red line and set a filter on Harry station.

Step 5 (11:07) participant navigated to the Yellow line in the Legend, listed the first 10 stations (to Lima), exploring Sierra and Lima, and set a filter on the Lima station.

The participant took a systematic approach to this task: visit each line in turn answer for the stations in question. The participant quickly found X-Ray station on the Blue line and set a filter on both. The

visit to the Green line turned out to be unfruitful. The visit to the Purple line turned up Zara station. The participant found Harry station next on the Red line and set a filter on both. This is the first indication that something has gone awry: the participant very quickly set filters on X-Ray, Harry, the Blue line and the Red line but when encountering Zara, merely parsed and then moved on. It seems that the participant was not sure whether that Zara was the one they were looking for. The participant next explored the Yellow line and made a point of exploring Sierra station, to which is on the Blue and Yellow lines. The participant then set a filter on Lima station. This course of action seems initially puzzling: why would Lima station be particularly important? However the participant has explored Sierra station. It could be that the participant thought Sierra station was in fact Zara station. The participant knows from the filters that Oscar station is a transfer point between the Red and Blue lines. Thus, discovering that Sierra station is on the Blue and Yellow lines, thinking that Sierra was Zara, offers the participant a tidy solution. This is shown in Figure D.5.

Appendix E

ANOVA tables

E.1 Task performance

Total time

```
Analysis of Variance for log of total time, using Adjusted SS for Tests

Source  DF   Seq SS   Adj SS   Adj MS     F      P
cond     1  0.24552  0.24821  0.24821  7.37  0.008
task     3  0.96384  0.96384  0.32128  9.54  0.000
Error   76  2.56015  2.56015  0.03369
Total   80  3.76951
```

```
Least Squares Means for log of total time

cond        Mean  SE Mean
overview   3.063  0.02838
speech     2.952  0.02952
```

```
         Difference        SE of              Adjusted
cond      of Means    Difference   T-Value    P-Value
speech     -0.1109       0.04085    -2.714     0.0082
```

Number of moves

```
Analysis of Variance for num moves, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS   Adj MS     F      P
cond     1     14046     14452    14452  0.54  0.466
task     3    460252    460252   153417  5.69  0.001
Error   76   2049798   2049798    26971
Total   80   2524097
```

```
Least Squares Means for num moves

cond        Mean   SE Mean
overview   362.6     25.39
speech     389.3     26.42



        Difference       SE of           Adjusted
cond      of Means  Difference  T-Value   P-Value
speech       26.76       36.55   0.7320    0.4664
```

Average move time

```
Analysis of Variance for log of Average move time, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS      F      P
cond     1  0.316220   0.322991   0.322991  40.21  0.000
task     3  0.069030   0.069030   0.023010   2.86  0.042
Error   76  0.610541   0.610541   0.008033
Total   80  0.995791



Least Squares Means for log of Average move time

cond        Mean   SE Mean
overview  0.6670   0.01386
speech    0.5405   0.01442



        Difference       SE of           Adjusted
cond      of Means  Difference  T-Value   P-Value
speech     -0.1265     0.01995   -6.341    0.0000
```

Score

```
Analysis of Variance for score, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS   Adj MS      F      P
cond     1     228.3      204.0    204.0   0.31  0.576
task     3     535.0      535.0    178.3   0.28  0.843
Error   76   49212.7    49212.7    647.5
Total   80   49976.0



Least Squares Means for score

cond        Mean   SE Mean
overview   67.55     3.934
speech     70.73     4.093
```

```
        Difference        SE of            Adjusted
cond      of Means  Difference  T-Value    P-Value
speech       3.179       5.664   0.5612     0.5763
```

E.2 Legend vs. Index usage

Count index

```
Analysis of Variance for count index, using Adjusted SS for Tests

Source  DF   Seq SS    Adj SS  Adj MS     F      P
cond     1    8.791     8.892   8.892  6.65  0.012
task     3    7.639     7.639   2.546  1.91  0.136
Error   76  101.570   101.570   1.336
Total   80  118.000
```

```
Least Squares Means for count index

cond        Mean  SE Mean
overview  2.092   0.1787
speech    1.428   0.1859
```

```
        Difference        SE of            Adjusted
cond      of Means  Difference  T-Value    P-Value
speech     -0.6637      0.2573   -2.579     0.0118
```

Count legend

```
Analysis of Variance for count legend, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS  Adj MS     F      P
cond     1     2.741     2.674   2.674  1.56  0.215
task     3    12.229    12.229   4.076  2.38  0.076
Error   76   129.919   129.919   1.709
Total   80   144.889
```

```
Least Squares Means for count legend

cond        Mean  SE Mean
overview  1.881   0.2021
speech    1.517   0.2103
```

```
        Difference        SE of            Adjusted
cond      of Means  Difference  T-Value    P-Value
speech     -0.3640      0.2910   -1.251     0.2149
```

Context changes

```
Analysis of Variance for context changes, using Adjusted SS for Tests

Source  DF   Seq SS    Adj SS   Adj MS     F      P
cond     1   21.349    21.318   21.318   3.71   0.058
task     3   38.352    38.352   12.784   2.23   0.092
Error   76  436.522   436.522    5.744
Total   80  496.222
```

```
Least Squares Means for context changes

cond        Mean  SE Mean
overview   2.973   0.3705
speech     1.945   0.3855
```

```
        Difference       SE of              Adjusted
cond      of Means   Difference   T-Value    P-Value
speech      -1.028       0.5334    -1.927     0.0578
```

Time legend

```
Analysis of Variance for time legend, using Adjusted SS for Tests

Source  DF   Seq SS    Adj SS   Adj MS     F      P
cond     1  21237.1   21595.0  21595.0  33.53   0.000
task     3   1296.0    1296.0    432.0   0.67   0.573
Error   76  48954.2   48954.2    644.1
Total   80  71487.3
```

```
Least Squares Means for time legend

cond        Mean  SE Mean
overview   37.72    3.924
speech     70.43    4.082
```

```
        Difference       SE of              Adjusted
cond      of Means   Difference   T-Value    P-Value
speech       32.71        5.649     5.790     0.0000
```

Time index

```
Analysis of Variance for time index, using Adjusted SS for Tests

Source  DF   Seq SS    Adj SS   Adj MS     F      P
cond     1  21164.2   21522.9  21522.9  33.42   0.000
task     3   1306.5    1306.5    435.5   0.68   0.569
```

```
Error    76  48948.3  48948.3    644.1
Total    80  71418.9
```

Least Squares Means for time index

```
cond        Mean  SE Mean
overview   62.22    3.924
speech     29.57    4.082
```

```
        Difference      SE of          Adjusted
cond      of Means  Difference  T-Value    P-Value
speech      -32.65       5.649   -5.781     0.0000
```

Moves legend

Analysis of Variance for moves legend, using Adjusted SS for Tests

```
Source  DF   Seq SS   Adj SS   Adj MS      F      P
cond     1  17381.8  17624.1  17624.1  27.88  0.000
task     3    777.7    777.7    259.2   0.41  0.746
Error   76  48044.1  48044.1    632.2
Total   80  66203.6
```

Least Squares Means for moves legend

```
cond        Mean  SE Mean
overview   34.48    3.887
speech     64.02    4.044
```

```
        Difference      SE of          Adjusted
cond      of Means  Difference  T-Value    P-Value
speech       29.55       5.596    5.280     0.0000
```

Moves index

Analysis of Variance for moves index, using Adjusted SS for Tests

```
Source  DF   Seq SS   Adj SS   Adj MS      F       P
cond     1  17381.8  17624.1  17624.1  27.88  0.000
task     3    777.7    777.7    259.2   0.41  0.746
Error   76  48044.1  48044.1    632.2
Total   80  66203.6
```

Least Squares Means for moves index

```
cond       Mean  SE Mean
overview  65.52    3.887
speech    35.98    4.044
```

```
        Difference      SE of          Adjusted
cond     of Means  Difference  T-Value   P-Value
speech     -29.55       5.596   -5.280    0.0000
```

Moves per turn, legend

```
Analysis of Variance for mpt legend (n), using Adjusted SS for Tests

Source  DF   Seq SS   Adj SS   Adj MS      F      P
cond     1  22565.9  22852.8  22852.8  37.17  0.000
task     3    872.3    872.3    290.8   0.47  0.702
Error   76  46728.1  46728.1    614.8
Total   80  70166.3
```

```
Least Squares Means for mpt legend (n)
```

```
cond       Mean  SE Mean
overview  19.25    3.834
speech    52.90    3.988
```

```
        Difference      SE of          Adjusted
cond     of Means  Difference  T-Value   P-Value
speech      33.65       5.519    6.097    0.0000
```

Moves per turn, index

F-value ranked 91 out of 10,000 in Monte Carlo Simulation (MCS)

```
Analysis of Variance for mpt index (n), using Adjusted SS for Tests

Source  DF   Seq SS   Adj SS   Adj MS     F      P
cond     1   6600.6   6758.3   6758.3  8.62  0.004
task     3   2896.7   2896.7    965.6  1.23  0.304
Error   76  59566.4  59566.4    783.8
Total   80  69063.7
```

```
Least Squares Means for mpt index (n)
```

```
cond       Mean  SE Mean
overview  45.02    4.328
speech    26.72    4.503
```

```
        Difference       SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech       -18.30       6.231   -2.936    0.0044
```

Time per turn, legend

```
Analysis of Variance for tpt legend (n), using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS    Adj MS       F      P
cond     1   27476.6   27850.0   27850.0   44.89  0.000
task     3     936.8     936.8     312.3    0.50  0.681
Error   76   47151.5   47151.5     620.4
Total   80   75564.9
```

```
Least Squares Means for tpt legend (n)

cond        Mean   SE Mean
overview   20.44     3.851
speech     57.58     4.006
```

```
        Difference       SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech        37.14       5.544    6.700    0.0000
```

Time per turn, index (n)

```
Analysis of Variance for tpt index (n), using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS   Adj MS       F      P
cond     1    9036.1    9248.0   9248.0   11.53  0.001
task     3    2912.3    2912.3    970.8    1.21  0.312
Error   76   60969.7   60969.7    802.2
Total   80   72918.1
```

```
Least Squares Means for tpt index (n)

cond        Mean   SE Mean
overview   43.40     4.379
speech     22.00     4.556
```

```
        Difference       SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech       -21.40       6.304   -3.395    0.0011
```

E.3   Usage at each level

Level 3 moves

```
Analysis of Variance for level 3 moves, using Adjusted SS for Tests


Source  DF   Seq SS   Adj SS   Adj MS      F       P
cond     1  0.60268  0.60645  0.60645  35.53  0.000
task     3  0.08270  0.08270  0.02757   1.61  0.193
Error   76  1.29735  1.29735  0.01707
Total   80  1.98272



Least Squares Means for level 3 moves


cond          Mean   SE Mean
overview  0.4395   0.02020
speech    0.2661   0.02102



        Difference        SE of              Adjusted
cond      of Means   Difference  T-Value    P-Value
speech     -0.1733      0.02908   -5.960     0.0000
```

Level 4 moves

Ranked 10 out of 10,000 in MCS.

```
Analysis of Variance for level 4 moves, using Adjusted SS for Tests


Source  DF    Seq SS     Adj SS     Adj MS      F       P
cond     1  0.061244  0.060326  0.060326  22.51  0.000
task     3  0.008644  0.008644  0.002881   1.08  0.365
Error   76  0.203645  0.203645  0.002680
Total   80  0.273533



Least Squares Means for level 4 moves


cond          Mean    SE Mean
overview  0.14580   0.008003
speech    0.09113   0.008326



        Difference        SE of              Adjusted
cond      of Means   Difference  T-Value    P-Value
speech    -0.05467      0.01152   -4.745     0.0000
```

Level 5 moves

```
Analysis of Variance for level 5 moves, using Adjusted SS for Tests
```

```
Source   DF   Seq SS    Adj SS    Adj MS      F       P
cond      1  0.46562   0.46853   0.46853  29.69   0.000
task      3  0.01147   0.01147   0.00382   0.24   0.867
Error    76  1.19929   1.19929   0.01578
Total    80  1.67638
```

```
Least Squares Means for level 5 moves
```

```
cond          Mean  SE Mean
overview  0.2490   0.01942
speech    0.4013   0.02021
```

```
         Difference       SE of               Adjusted
cond       of Means   Difference  T-Value    P-Value
speech       0.1524      0.02796    5.449      0.0000
```

Level 6 moves

```
Analysis of Variance for level 6 moves, using Adjusted SS for Tests
```

```
Source   DF    Seq SS     Adj SS     Adj MS      F      P
cond      1  0.011411   0.010953   0.010953   5.54   0.021
task      3  0.006720   0.006720   0.002240   1.13   0.341
Error    76  0.150347   0.150347   0.001978
Total    80  0.168478
```

```
Least Squares Means for level 6 moves
```

```
cond           Mean   SE Mean
overview  0.06516   0.006877
speech    0.08846   0.007154
```

```
         Difference       SE of               Adjusted
cond       of Means   Difference  T-Value    P-Value
speech      0.02329     0.009900    2.353      0.0212
```

Level 7 moves

Ranked 98 out of 10,000 in MCS.

```
Analysis of Variance for level 7 moves, using Adjusted SS for Tests
```

```
Source   DF    Seq SS     Adj SS     Adj MS      F       P
cond      1  0.093080   0.076217   0.076217  22.90   0.000
exp      21  0.299443   0.299702   0.014272   4.29   0.000
```

```
task     3  0.001782  0.001782  0.000594    0.18  0.911
Error   55  0.183046  0.183046  0.003328
Total   80  0.577350
```

Least Squares Means for level 7 moves

```
cond         Mean   SE Mean
overview  0.05265  0.008995
speech    0.11624  0.009805
```

```
         Difference       SE of              Adjusted
cond       of Means  Difference  T-Value    P-Value
speech      0.06359     0.01329    4.786     0.0000
```

Count unique all

Analysis of Variance for log of Count unique all, using Adjusted SS for Tests

```
Source  DF   Seq SS   Adj SS   Adj MS      F       P
cond     1  0.19501  0.20160  0.20160  13.51   0.000
task     3  0.33539  0.33539  0.11180   7.49   0.000
Error   76  1.13371  1.13371  0.01492
Total   80  1.66411
```

Least Squares Means for log of Count unique all

```
cond        Mean  SE Mean
overview   2.089  0.01888
speech     2.189  0.01965
```

```
         Difference       SE of              Adjusted
cond       of Means  Difference  T-Value    P-Value
speech      0.09994     0.02718    3.676     0.0004
```

Count unique 3

Analysis of Variance for count unique 3, using Adjusted SS for Tests

```
Source  DF    Seq SS   Adj SS   Adj MS      F      P
cond     1      20.4     24.5     24.5   0.08  0.772
task     3    8229.2   8229.2   2743.1   9.44  0.000
Error   76   22080.4  22080.4    290.5
Total   80   30330.0
```

Least Squares Means for count unique 3

```
cond        Mean   SE Mean
overview  52.71    2.635
speech    53.81    2.742
```

```
          Difference      SE of              Adjusted
cond        of Means  Difference  T-Value   P-Value
speech         1.102       3.794   0.2905    0.7723
```

Count unique 4

Analysis of Variance for log Count unique 4, using Adjusted SS for Tests

```
Source  DF   Seq SS    Adj SS   Adj MS      F       P
cond     1  0.44544   0.43362  0.43362  11.42   0.001
task     3  0.33801   0.33801  0.11267   2.97   0.037
Error   76  2.88548   2.88548  0.03797
Total   80  3.66893
```

Least Squares Means for log Count unique 4

```
cond        Mean  SE Mean
overview  1.226  0.03013
speech    1.079  0.03134
```

```
          Difference      SE of              Adjusted
cond        of Means  Difference  T-Value   P-Value
speech       -0.1466     0.04337   -3.380    0.0012
```

Count unique 5

Analysis of Variance for count unique 5, using Adjusted SS for Tests

```
Source  DF    Seq SS   Adj SS   Adj MS      F      P
cond     1    4366.3   4435.2   4435.2   9.12  0.003
task     3    3047.6   3047.6   1015.9   2.09  0.109
Error   76   36956.7  36956.7    486.3
Total   80   44370.5
```

Least Squares Means for count unique 5

```
cond        Mean   SE Mean
overview  33.53    3.409
speech    48.35    3.547
```

```
        Difference       SE of              Adjusted
cond      of Means  Difference  T-Value    P-Value
speech       14.82       4.908    3.020     0.0035
```

Count unique 6

```
Analysis of Variance for count unique 6, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS   Adj MS      F      P
cond     1    376.21    369.42   369.42   6.23   0.015
task     3    374.68    374.68   124.89   2.11   0.106
Error   76   4503.34   4503.34    59.25
Total   80   5254.22
```

```
Least Squares Means for count unique 6

cond          Mean   SE Mean
overview   10.96      1.190
speech     15.24      1.238
```

```
        Difference       SE of              Adjusted
cond      of Means  Difference  T-Value    P-Value
speech       4.278       1.713    2.497     0.0147
```

Count unique 7

```
Analysis of Variance for log of Count unique 7, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS   Adj MS       F       P
cond     1    5.2098    5.1840   5.1840   14.08   0.000
task     3    0.7470    0.7470   0.2490    0.68   0.569
Error   76   27.9757   27.9757   0.3681
Total   80   33.9325
```

```
Least Squares Means for log of Count unique 7

cond          Mean   SE Mean
overview   0.6688   0.09380
speech     1.1755   0.09759
```

```
        Difference       SE of              Adjusted
cond      of Means  Difference  T-Value    P-Value
speech      0.5068      0.1350    3.753     0.0003
```

Level 3 time

```
Analysis of Variance for level 3 time, using Adjusted SS for Tests
```

```
Source  DF    Seq SS     Adj SS     Adj MS      F       P
cond     1  0.070760   0.070175   0.070175   10.47   0.002
task     3  0.027770   0.027770   0.009257    1.38   0.255
Error   76  0.509414   0.509414   0.006703
Total   80  0.607945
```

Least Squares Means for level 3 time

```
cond         Mean  SE Mean
overview  0.2521  0.01266
speech    0.1931  0.01317
```

```
        Difference        SE of           Adjusted
cond     of Means  Difference  T-Value   P-Value
speech   -0.05896     0.01822   -3.236    0.0018
```

Level 4 time

Analysis of Variance for level 4 time, using Adjusted SS for Tests

```
Source  DF   Seq SS   Adj SS   Adj MS      F      P
cond     1  0.48800  0.48432  0.48432  65.02  0.000
task     3  0.05298  0.05298  0.01766   2.37  0.077
Error   76  0.56614  0.56614  0.00745
Total   80  1.10712
```

Least Squares Means for level 4 time

```
cond          Mean  SE Mean
overview  0.24996  0.01334
speech    0.09506  0.01388
```

```
        Difference        SE of           Adjusted
cond     of Means  Difference  T-Value   P-Value
speech    -0.1549     0.01921   -8.063    0.0000
```

Level 5 time

Analysis of Variance for level 5 time, using Adjusted SS for Tests

```
Source  DF   Seq SS   Adj SS   Adj MS      F      P
cond     1  1.00984  1.01147  1.01147  71.78  0.000
task     3  0.02804  0.02804  0.00935   0.66  0.577
Error   76  1.07095  1.07095  0.01409
Total   80  2.10883
```

```
Least Squares Means for level 5 time

cond         Mean   SE Mean
overview   0.2381   0.01835
speech     0.4619   0.01909


         Difference       SE of            Adjusted
cond      of Means   Difference   T-Value   P-Value
speech      0.2239      0.02642     8.472    0.0000
```

Level 6 time

Ranked 68 out of 10,000 in MCS.

```
Analysis of Variance for level 6 time, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS      F      P
cond     1  0.087159   0.090344   0.090344   9.97   0.002
task     3  0.027557   0.027557   0.009186   1.01   0.392
Error   76  0.689002   0.689002   0.009066
Total   80  0.803718


Least Squares Means for level 6 time

cond          Mean   SE Mean
overview   0.15848   0.01472
speech     0.09158   0.01531


         Difference       SE of            Adjusted
cond      of Means   Difference   T-Value   P-Value
speech     -0.06690      0.02119    -3.157    0.0023
```

Level 7 time

Ranked 156 out of 10,000 in MCS.

```
Analysis of Variance for level 7 time, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS      F      P
cond     1  0.104850   0.104830   0.104830  17.08   0.000
task     3  0.000556   0.000556   0.000185   0.03   0.993
Error   76  0.466333   0.466333   0.006136
Total   80  0.571738
```

```
Least Squares Means for level 7 time

cond          Mean   SE Mean
overview   0.05166   0.01211
speech     0.12372   0.01260



         Difference       SE of          Adjusted
cond      of Means   Difference   T-Value   P-Value
speech     0.07207      0.01744     4.133    0.0001
```

E.4 Level usage in the Legend

Legend level 3 time

Ranked 884 out of 10,000 in MCS. Not significant.

```
Analysis of Variance for legend L3 time, using Adjusted SS for Tests

Source  DF      Seq SS       Adj SS       Adj MS      F      P
cond     1   0.0033649    0.0036089    0.0036089   4.34   0.041
task     3   0.0106781    0.0106781    0.0035594   4.28   0.008
Error   76   0.0631684    0.0631684    0.0008312
Total   80   0.0772114



Least Squares Means for legend L3 time

cond          Mean   SE Mean
overview   0.04196   0.004457
speech     0.05533   0.004637



         Difference       SE of          Adjusted
cond      of Means   Difference   T-Value   P-Value
speech     0.01337     0.006417     2.084    0.0406
```

Legend level 3 moves

```
Analysis of Variance for legend level 3 moves, using Adjusted SS for Tests

Source  DF     Seq SS      Adj SS      Adj MS      F      P
cond     1   0.000065    0.000033    0.000033   0.02   0.877
task     3   0.003803    0.003803    0.001268   0.93   0.429
Error   76   0.103228    0.103228    0.001358
Total   80   0.107096



Least Squares Means for legend level 3 moves
```

```
cond        Mean    SE Mean
overview  0.05799  0.005698
speech    0.05672  0.005928
```

```
        Difference       SE of            Adjusted
cond      of Means  Difference  T-Value   P-Value
speech   -0.001277    0.008203  -0.1557    0.8767
```

Legend level 3 unique

Ranked 25 out of 10,000 in MCS.

```
Analysis of Variance for legend L3 unique, using Adjusted SS for Tests

Source  DF   Seq SS   Adj SS  Adj MS      F      P
cond     1   40.218   42.003  42.003  17.07  0.000
task     3   25.759   25.759   8.586   3.49  0.020
Error   76  187.011  187.011   2.461
Total   80  252.988
```

```
Least Squares Means for legend L3 unique

cond       Mean  SE Mean
overview  3.343   0.2425
speech    4.786   0.2523
```

```
        Difference       SE of            Adjusted
cond      of Means  Difference  T-Value   P-Value
speech       1.443      0.3491    4.132    0.0001
```

Ranked 3908 out of 10,000 in MCS. Not significant.

Legend level 4 time

```
Analysis of Variance for legend L4 time, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS   Adj MS     F      P
cond     1  0.022147  0.020352 0.020352  4.59  0.035
task     3  0.024346  0.024346 0.008115  1.83  0.149
Error   76  0.336810  0.336810 0.004432
Total   80  0.383303
```

```
Least Squares Means for legend L4 time

cond        Mean  SE Mean
overview  0.09294  0.01029
```

```
speech    0.06119  0.01071
```

```
        Difference      SE of           Adjusted
cond      of Means  Difference  T-Value   P-Value
speech    -0.03175     0.01482   -2.143    0.0353
```

Legend level 4 moves

```
Analysis of Variance for legend L4 moves, using Adjusted SS for Tests

Source  DF     Seq SS     Adj SS     Adj MS     F       P
cond     1   0.000009   0.000000   0.000000  0.00   0.995
task     3   0.004742   0.004742   0.001581  1.23   0.304
Error   76   0.097466   0.097466   0.001282
Total   80   0.102217
```

```
Least Squares Means for legend L4 moves

cond          Mean    SE Mean
overview   0.05361   0.005537
speech     0.05355   0.005760
```

```
        Difference        SE of              Adjusted
cond      of Means    Difference   T-Value    P-Value
speech   -0.000053     0.007971  -0.006589     0.9948
```

Legend level 4 unique

```
Analysis of Variance for legend L4 unique, using Adjusted SS for Tests

Source  DF  Seq SS  Adj SS  Adj MS     F       P
cond     1    0.10    0.23    0.23  0.02   0.885
task     3   63.31   63.31   21.10  1.98   0.124
Error   76  809.95  809.95   10.66
Total   80  873.36
```

```
Least Squares Means for legend L4 unique

cond       Mean   SE Mean
overview  5.565    0.5047
speech    5.670    0.5251
```

```
        Difference        SE of           Adjusted
cond      of Means    Difference  T-Value   P-Value
speech      0.1058        0.7266   0.1456    0.8846
```

Legend level 5 time

```
Analysis of Variance for legend L5 time, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS    Adj MS       F       P
cond     1   1.38144   1.38855   1.38855   62.37   0.000
task     3   0.02569   0.02569   0.00856    0.38   0.764
Error   76   1.69207   1.69207   0.02226
Total   80   3.09920
```

```
Least Squares Means for legend L5 time

cond          Mean   SE Mean
overview    0.1480   0.02307
speech      0.4103   0.02400
```

```
         Difference        SE of           Adjusted
cond      of Means    Difference   T-Value    P-Value
speech      0.2623       0.03321     7.897     0.0000
```

Legend level 5 moves

```
Analysis of Variance for legend L5 moves, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS    Adj MS       F       P
cond     1   0.69073   0.69699   0.69699   28.71   0.000
task     3   0.01324   0.01324   0.00441    0.18   0.908
Error   76   1.84474   1.84474   0.02427
Total   80   2.54872
```

```
Least Squares Means for legend L5 moves

cond          Mean   SE Mean
overview    0.1743   0.02409
speech      0.3601   0.02506
```

```
         Difference        SE of           Adjusted
cond      of Means    Difference   T-Value    P-Value
speech      0.1858       0.03468     5.359     0.0000
```

Legend level 5 unique

```
Analysis of Variance for legend L5 unique, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS   Adj MS       F      P
cond     1    4584.5    4619.3   4619.3   11.92   0.001
task     3    1825.1    1825.1    608.4    1.57   0.204
```

```
Error   76  29461.4  29461.4   387.6
Total   80  35871.0
```

Least Squares Means for legend L5 unique

```
cond        Mean  SE Mean
overview  26.63    3.044
speech    41.75    3.167
```

```
         Difference      SE of           Adjusted
cond      of Means   Difference  T-Value   P-Value
speech      15.13        4.382    3.452    0.0009
```

Legend level 6 time

Analysis of Variance for legend L6 time, using Adjusted SS for Tests

```
Source  DF    Seq SS     Adj SS     Adj MS     F      P
cond     1  0.014940  0.015609  0.015609  3.61  0.061
task     3  0.011042  0.011042  0.003681  0.85  0.471
Error   76  0.328976  0.328976  0.004329
Total   80  0.354958
```

Least Squares Means for legend L6 time

```
cond        Mean  SE Mean
overview  0.04958  0.01017
speech    0.07738  0.01058
```

```
         Difference      SE of           Adjusted
cond      of Means   Difference  T-Value   P-Value
speech     0.02781      0.01464    1.899    0.0614
```

Legend level 6 moves

Ranked 228 out of 10,000 in MCS.

Analysis of Variance for legend L6 moves, using Adjusted SS for Tests

```
Source  DF    Seq SS     Adj SS     Adj MS     F      P
cond     1  0.048233  0.048071  0.048071  23.47  0.000
task     3  0.001175  0.001175  0.000392   0.19  0.902
Error   76  0.155685  0.155685  0.002048
Total   80  0.205093
```

```
Least Squares Means for legend L6 moves

cond          Mean    SE Mean
overview   0.02667   0.006998
speech     0.07548   0.007280



         Difference       SE of           Adjusted
cond        of Means  Difference  T-Value   P-Value
speech       0.04880     0.01007    4.844    0.0000
```

Legend level 6 unique

```
Analysis of Variance for legend L6 unique, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS    Adj MS       F       P
cond     1   1186.35   1190.86   1190.86   16.96   0.000
task     3    174.36    174.36     58.12    0.83   0.483
Error   76   5334.85   5334.85     70.20
Total   80   6695.56



Least Squares Means for legend L6 unique

cond         Mean   SE Mean
overview    5.194     1.295
speech     12.875     1.348



         Difference       SE of           Adjusted
cond        of Means  Difference  T-Value   P-Value
speech         7.681       1.865    4.119    0.0001
```

Legend level 7 time

Ranked 227 out of 10,000 in MCS.

```
Analysis of Variance for legend L7 time, using Adjusted SS for Tests

Source  DF     Seq SS     Adj SS     Adj MS       F       P
cond     1   0.079908   0.079047   0.079047   31.79   0.000
task     3   0.001516   0.001516   0.000505    0.20   0.894
Error   76   0.189004   0.189004   0.002487
Total   80   0.270428



Least Squares Means for legend L7 time

cond          Mean    SE Mean
overview   0.01981   0.007710
```

```
speech    0.08239  0.008021
```

```
         Difference        SE of            Adjusted
cond        of Means  Difference  T-Value  P-Value
speech       0.06258     0.01110    5.638   0.0000
```

Legend level 7 moves

Ranked 243 out of 10,000 in MCS.

```
Analysis of Variance for legend L7 moves, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS      F       P
cond     1  0.063832   0.063377   0.063377  30.50   0.000
task     3  0.000432   0.000432   0.000144   0.07   0.976
Error   76  0.157899   0.157899   0.002078
Total   80  0.222164
```

```
Least Squares Means for legend L7 moves

cond         Mean    SE Mean
overview  0.01861   0.007047
speech    0.07464   0.007332
```

```
         Difference        SE of            Adjusted
cond        of Means  Difference  T-Value  P-Value
speech       0.05603     0.01015    5.523   0.0000
```

Legend level 7 unique

Ranked 139 out of 10,000 in MCS.

```
Analysis of Variance for legend L7 unique, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS    Adj MS      F       P
cond     1  1984.62   1979.83   1979.83  30.69   0.000
task     3    46.99     46.99     15.66   0.24   0.866
Error   76  4902.39   4902.39     64.51
Total   80  6934.00
```

```
Least Squares Means for legend L7 unique

cond        Mean   SE Mean
overview   3.134    1.242
speech    13.038    1.292
```

```
        Difference        SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech        9.904       1.788    5.540    0.0000
```

E.5   Level usage in the index

Index level 3 time

```
Analysis of Variance for index L3 time, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS      F       P
cond     1  0.104987   0.105612   0.105612  13.68   0.000
task     3  0.027675   0.027675   0.009225   1.20   0.317
Error   76  0.586640   0.586640   0.007719
Total   80  0.719302
```

```
Least Squares Means for index L3 time

cond         Mean   SE Mean
overview  0.2101   0.01358
speech    0.1378   0.01413
```

```
        Difference        SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech     -0.07233     0.01956   -3.699    0.0004
```

Index level 3 moves

```
Analysis of Variance for index L3 moves, using Adjusted SS for Tests

Source  DF    Seq SS    Adj SS    Adj MS      F       P
cond     1   0.63578   0.64602   0.64602  31.64   0.000
task     3   0.10829   0.10829   0.03610   1.77   0.160
Error   76   1.55171   1.55171   0.02042
Total   80   2.29577
```

```
Least Squares Means for index L3 moves

cond         Mean   SE Mean
overview  0.3847   0.02209
speech    0.2058   0.02298
```

```
        Difference        SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech      -0.1789     0.03180   -5.625    0.0000
```

Index level 3 unique

```
Analysis of Variance for index L3 unique, using Adjusted SS for Tests

Source  DF   Seq SS    Adj SS   Adj MS      F      P
cond     1      3.8       2.8      2.8   0.01  0.922
task     3   7694.1    7694.1   2564.7   8.84  0.000
Error   76  22038.3   22038.3    290.0
Total   80  29736.2


Least Squares Means for index L3 unique

cond         Mean  SE Mean
overview    49.32    2.633
speech      48.95    2.739


         Difference         SE of                Adjusted
cond       of Means    Difference   T-Value     P-Value
speech      -0.3727         3.790  -0.09834      0.9219
```

Index level 4 time

```
Analysis of Variance for index L4 time, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS    Adj MS       F      P
cond     1  0.302227   0.306106  0.306106   55.51  0.000
task     3  0.034051   0.034051  0.011350    2.06  0.113
Error   76  0.419126   0.419126  0.005515
Total   80  0.755404


Least Squares Means for index L4 time

cond          Mean  SE Mean
overview   0.15702  0.01148
speech     0.03388  0.01194


         Difference         SE of                Adjusted
cond       of Means    Difference   T-Value     P-Value
speech      -0.1231       0.01653    -7.450      0.0000
```

Index level 4 moves

Ranked 23 out of 10,000 in MCS.

```
Analysis of Variance for index L4 moves, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS    Adj MS       F         P
```

```
cond      1  0.059776  0.060210  0.060210  21.25  0.000
task      3  0.004295  0.004295  0.001432   0.51  0.680
Error    76  0.215367  0.215367  0.002834
Total    80  0.279439
```

```
Least Squares Means for index L4 moves
```

```
cond           Mean    SE Mean
overview   0.09219   0.008230
speech     0.03758   0.008562
```

```
         Difference       SE of              Adjusted
cond       of Means  Difference  T-Value   P-Value
speech     -0.05462     0.01185   -4.609    0.0000
```

Index level 4 unique

```
Analysis of Variance for index L4 unique, using Adjusted SS for Tests
```

```
Source  DF    Seq SS   Adj SS  Adj MS     F      P
cond     1     441.4    429.9   429.9  3.11  0.082
task     3     486.6    486.6   162.2  1.17  0.326
Error   76   10505.0  10505.0   138.2
Total   80   11433.0
```

```
Least Squares Means for index L4 unique
```

```
cond         Mean  SE Mean
overview   12.167    1.818
speech      7.552    1.891
```

```
         Difference       SE of              Adjusted
cond       of Means  Difference  T-Value   P-Value
speech       -4.615       2.617   -1.763    0.0818
```

Index level 5 time

Ranked 143 out of 10,000 in MCS.

```
Analysis of Variance for index L5 time, using Adjusted SS for Tests
```

```
Source  DF    Seq SS    Adj SS    Adj MS     F      P
cond     1  0.029050  0.029808  0.029808  8.68  0.004
task     3  0.007143  0.007143  0.002381  0.69  0.559
Error   76  0.261078  0.261078  0.003435
Total   80  0.297271
```

```
Least Squares Means for index L5 time

cond          Mean     SE Mean
overview   0.09010   0.009062
speech     0.05167   0.009427


          Difference        SE of              Adjusted
cond       of Means    Difference   T-Value    P-Value
speech      -0.03843      0.01305    -2.946     0.0043
```

Index level 5 moves

Ranked 163 out of 10,000 in MCS.

```
Analysis of Variance for index L5 moves, using Adjusted SS for Tests

Source  DF     Seq SS     Adj SS     Adj MS     F        P
cond     1   0.022124   0.022610   0.022610   8.86   0.004
task     3   0.005637   0.005637   0.001879   0.74   0.534
Error   76   0.193941   0.193941   0.002552
Total   80   0.221702


Least Squares Means for index L5 moves

cond          Mean     SE Mean
overview   0.07467   0.007810
speech     0.04120   0.008125


          Difference        SE of              Adjusted
cond       of Means    Difference   T-Value    P-Value
speech      -0.03347      0.01124    -2.977     0.0039
```

Index level 5 unique

```
Analysis of Variance for index L5 unique, using Adjusted SS for Tests

Source  DF  Seq SS  Adj SS  Adj MS     F       P
cond     1     2.7     1.9     1.9   0.01   0.904
task     3   177.1   177.1    59.0   0.46   0.709
Error   76  9692.2  9692.2   127.5
Total   80  9872.0


Least Squares Means for index L5 unique
```

```
cond       Mean   SE Mean
overview  6.902    1.746
speech    6.598    1.816
```

```
          Difference        SE of              Adjusted
cond        of Means   Difference  T-Value    P-Value
speech       -0.3045        2.514  -0.1212     0.9039
```

Index level 6 time

Ranked 6 out of 10,000 in MCS.

```
Analysis of Variance for index L6 time, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS      F      P
cond     1  0.174271   0.181057   0.181057  19.43  0.000
task     3  0.040996   0.040996   0.013665   1.47  0.230
Error   76  0.708297   0.708297   0.009320
Total   80  0.923564
```

```
Least Squares Means for index L6 time

cond        Mean   SE Mean
overview  0.10891   0.01493
speech    0.01420   0.01553
```

```
          Difference        SE of              Adjusted
cond        of Means   Difference  T-Value    P-Value
speech      -0.09471      0.02149   -4.408     0.0000
```

Index level 6 moves

Ranked 34 out of 10,000 in MCS.

```
Analysis of Variance for index L6 moves, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS      F      P
cond     1  0.012723   0.013132   0.013132  12.34  0.001
task     3  0.003899   0.003899   0.001300   1.22  0.308
Error   76  0.080847   0.080847   0.001064
Total   80  0.097469
```

```
Least Squares Means for index L6 moves

cond        Mean    SE Mean
overview  0.03849   0.005043
```

```
speech    0.01298  0.005246
```

```
        Difference      SE of             Adjusted
cond      of Means  Difference  T-Value   P-Value
speech     -0.02551   0.007260   -3.514    0.0008
```

Index level 6 unique

Ranked 33 out of 10,000 in MCS.

```
Analysis of Variance for index L6 unique, using Adjusted SS for Tests

Source  DF   Seq SS   Adj SS  Adj MS      F      P
cond     1   226.42   233.74  233.74  15.82  0.000
task     3    78.77    78.77   26.26   1.78  0.159
Error   76  1122.81  1122.81   14.77
Total   80  1428.00
```

```
Least Squares Means for index L6 unique

cond       Mean  SE Mean
overview  5.769   0.5943
speech    2.366   0.6182
```

```
        Difference      SE of             Adjusted
cond      of Means  Difference  T-Value   P-Value
speech      -3.403      0.8555   -3.978    0.0002
```

Index level 7 time

```
Analysis of Variance for index L7 time, using Adjusted SS for Tests

Source  DF    Seq SS     Adj SS     Adj MS     F      P
cond     1  0.001691   0.001817   0.001817  0.41  0.523
task     3  0.001522   0.001522   0.000507  0.11  0.951
Error   76  0.335917   0.335917   0.004420
Total   80  0.339131
```

```
Least Squares Means for index L7 time

cond         Mean  SE Mean
overview  0.03185  0.01028
speech    0.04133  0.01069
```

```
        Difference      SE of             Adjusted
```

```
cond       of Means  Difference  T-Value   P-Value
speech    0.009487     0.01480    0.6411    0.5234
```

Index level 7 moves

Analysis of Variance for index L7 moves, using Adjusted SS for Tests

```
Source  DF    Seq SS     Adj SS     Adj MS     F      P
cond     1  0.002750   0.002910   0.002910  0.52  0.472
task     3  0.002944   0.002944   0.000981  0.18  0.912
Error   76  0.423752   0.423752   0.005576
Total   80  0.429446
```

Least Squares Means for index L7 moves

```
cond          Mean  SE Mean
overview  0.03335  0.01154
speech    0.04536  0.01201
```

```
        Difference       SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech      0.01201     0.01662   0.7224    0.4722
```

Index level 7 unique

Analysis of Variance for index L7 unique, using Adjusted SS for Tests

```
Source  DF   Seq SS    Adj SS   Adj MS     F      P
cond     1    237.7     249.9    249.9  0.75  0.388
task     3    433.2     433.2    144.4  0.43  0.729
Error   76  25233.1   25233.1    332.0
Total   80  25904.0
```

Least Squares Means for index L7 unique

```
cond          Mean  SE Mean
overview     8.304    2.817
speech      11.823    2.931
```

```
        Difference       SE of            Adjusted
cond       of Means  Difference  T-Value   P-Value
speech        3.519       4.056   0.8676    0.3883
```

Appendix F

Connect Four: a non-speech auditory overview

The overview that I describe in this evaluation was developed as part of the European COST action on Sonic Interaction Design (IC0601)[1]. The goal of the work was to investigate the sonification of grid-based games. One of the overviews generated as part of Sonic Interaction Design (SID) was Connect Four. Visually, the two dimensionality of the game grid can be quite powerful but is difficult to represent auditorily. In grid-based games, individual cell values are relatively meaningless: sets of cells together can represent a pattern of interest. As a result, the overview has to give a strong sense of context. These attributes make it a good choice for evaluating what the participants are able to understand from the overview.

F.1 The Connect Four game

Connect Four is a two-player game where the goal is to be the first to build up a line of four tokens. Each player has their own token and, in turn, places them in the 6×7 grid by dropping their tokens into columns. A winning line of four tokens can be vertical, horizontal or diagonal. It is a game that is both simple and sufficiently complex. Many will have encountered it as children. The rules are straight-forward and the grid is limited in size. It is these attributes that make it attractive as testbed for an auditory overview. Participants easily understand the rules and the scope of game play, so the focus is on the auditory representation instead. The goal was assess the auditory overview rather than participants recall. The audio represents a small grid of 42 cells. The number of cells, plus the layout of the two types of tokens in those cells means that participants will generally not be able to recall all the tokens entered in the grid simply from memory (Miller, 1956). The minimum number of tokens in play

---

[1] http://www.cost-sid.org/

	About the participants
Number	12
Sex	10 male, 2 female
Age	21-37 years (average 27)
Computer experience	11 expert users, 1 intermediate user
Musical experience	5 non-musicians, 7 with 5-16 years instrument playing (average 10)
Audio research	7 with no experience, 7 audio researchers
Connect Four experience	all

Table F.1: **Auditory Connect Four participant demographics.**

for a player to win is 7 – the 4 that make up the winning line and 3 of the opposite type from intervening turns. Informal tests during development of the interface indicated that players could keep track of 4-5 tokens (2-3 of each type) before misremembering token placement. As a result, the experiment described herein was a test of the auditory interface rather than participants memory of game play.

The description of the overview and what it sounds like is described in Nickerson & Hermann (2008). What follows are details on the evaluation subsequent to the publication of this paper.

F.2   Evaluating auditory Connect Four

In the experiment, pairs of participants were asked to play games against each other. After being trained on the various sounds in the overview and the graphics tablet (for token input), participants played three games, alternating which player went first.One and a half hours were allocated for each experiment and participants were paid £10 for their time. The games were then qualitatively analysed for evidence that the participants were actively using the audio to play rather than random chance.

F.2.1   About the participants

Six pairs of participants played a total of 18 auditory Connect Four games. They were recruited from the School of Electronic Engineering and Computer Science at Queen Mary University of London. I screened participants for musical ability, audio research and experience with Connect Four. The last was a requirement as I wanted to examine their use of the auditory overview. Table F.1 shows some demographics about the participants. Their ages ranged from 21 to 37 (average 27) and all were advanced computer users except one who considered himself intermediate. About half of them had musical experience and were involved in some way with audio research.

Code	Meaning
-2	bad move: player is clearly guessing and is unaware of the state of the grid, e.g. not blocking the opponent's winning move
-1	unwise move: player is potentially guessing
0	neutral move: no way to tell if the player is aware of the state of the grid or not, not a detrimental move
+1	good move: player shows evidence of strategy, is building a line of four tokens or blocks the opponent
+2	excellent move: player makes a clever move that wins the game or shows a long view of strategy

Table F.2: **The coding scheme for the analysis of auditory Connect Four.**

F.2.2 Game strategies

The games played in the experiment were an average of 14 moves long (min: 7, max: 33, median: 15, stdev: 4.5). As discussed above, it would be difficult to keep the entire game in working memory thus participants were relying on the audio to make their moves. Following from this assumption, any evidence of strategy would involve correct interpretation of the auditory overview.

In order to examine the games for strategy, I devised a coding scheme to rate each move. I found that 50% of moves (75% intercoder reliability) in the games demonstrated evidence of strategy, showing that the overview was sufficient to give a gross approximation of the state of the play grid during play.

Coding the games

The moves in the games were coded by me as well as two independent coders. The coding scheme is described in Table F.2. To facilitate the coding, I implemented an interface to graphically play back each game as well as step through them turn by turn. The procedure was as follows:

1. In preparation, coders played back each game to gain familiarity with what they would be coding

2. For each game, coders played back the game to get an overall impression

3. Having played the game back, coders would step through the game, rating each move as per the coding scheme

F.2.3 Participant feedback

After the games, participants provided both verbal and written feedback about their experience. The positive feedback centred on being able to successfully interpret the various auditory aspects of the overview, reinforcing the findings from analysing the strategy in the games. The negative feedback mostly related to the frustration of not being able to confirm what they thought they were hearing.

Auditory artefacts in support of understanding the overview

Participants commented on several aspects of the auditory design as useful in the playing of the game. Below I summarise the feedback where more than one of the participants commented. The attributes of the overview that these comments related to were focused on the scope of the overview, contextual information and the structure of the information in the grid. This relates back three of the characteristics of overviews as identified in Chapter 3.

Sonification of the columns/auditory timeline   Three participants commented that the representation of the columns worked well. This includes the drum beat used to indicate the start and middle of the grid. The use of a regular tempo was also beneficial. This indicates that the overview was successful at describing the scope of the game grid. (P3, exp 2; P5, exp 3; P8, exp 5)

Winning sound/adjacent tokens   Two participants commented on the sound of adjacent tokens. The brilliance of the token sounds was adjusted when there were several in a line. This indicated that the overview was successful at conveying important contextual information. (P4, exp 2; P6, exp 3)

Pitch indicating row   Two participants noted that the rising pitch of the token sounds to indicate the row it fell in was important in indicating how full a column was. This was a crucial part of understanding the structure of the information in the grid. (P2, exp 1;P5, exp 3)

What participants felt was missing

There were a number of traits of the overview that fell short of the participants' expectations. This can be largely explained by the fact that they were interacting with an overview and had no access to the detailed data. It can also be explained by the player's uncertainty about their interpretation of the sounds. Common comments were that the overview was too fast (P6, exp 3; P8, exp 4; P7, exp 4) and that it was hard to keep track of the opponent's tokens (P2, exp 1; P3, exp 2; P8, exp 4). Yet evidence from the games played indicated that they were able to interpret the sounds and make decisions based on them.

F.3   Summary

In this evaluation of auditory Connect Four, I was looking for evidence that participants could correctly interpret the information imparted by the overview and make use of it. This was the rationale behind looking for indications of strategy in the game play. That 50% of moves showed strategy confirms that without access to detailed data, the participants could use the audio or block and defeat their opponent.

The user feedback also confirmed that the information encoded in the overview was being interpreted. However, there was uncertainty on the part of the participants to trust what they were hearing. They were strategising yet desirous of slowing down the sonification to hear more clearly and found it challenging to understand their opponents' moves via the audio. This reinforces my belief that overviews need to be evaluated in conjunction with the detailed data so that the confidence users have in the overview can be increased. It was this finding that led to the format of my evaluation of the metro map overview in Chapter 4.

Bibliography

ALLMAN, T., DHILLON, R.K., LANDAU, M.A. & KURNIAWAN, S.H. (2009). Rock vibe: Rock band® computer games for people with no or limited vision. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*, 51–58, ACM, New York, NY, USA.

BARCHAS, J. (2003). *Graphic design, print culture, and the eighteenth-century novel*. Cambridge University Press, New York, NY, USA.

BARNICLE, K. (2000). Usability testing with screen reading technology in a windows environment. In *Proceedings of the Conference on Universal Usability*, 102–109, ACM, New York, NY, USA.

BARREAU, D. & NARDI, B.A. (1995). Finding and reminding: file organization from the desktop. *SIGCHI Bull.*, **27**, 39–43.

BEAUDOUIN-LAFON, M. & GAVER, W.W. (1994). Eno: synthesizing structured sound spaces. In *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*, 49–57, ACM, New York, NY, USA.

BERMAN, L.I. & GALLAGHER, K.B. (2009). Using sound to understand software architecture. In *Proceedings of the ACM International Conference on Design of Communication*, 127–134, ACM, New York, NY, USA.

BLATTNER, M.M., SUMIKAWA, D.A. & GREENBERG, R.M. (1989). Earcons and icons: their structure and common design principles. In *Human Computer Interaction*, vol. 4, 11–44.

BOVERMANN, T., HERMANN, T. & RITTER, H. (2006). Tangible data scanning sonification model. In T. Stockman, L.V. Nickerson, C. Frauenberger, A.D.N. Edwards & D. Brock, eds., *Proceedings of the International Conference on Auditory Display*, 77–82, Department of Computer Science, Queen Mary University of London, UK.

BREGMAN, A.S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press.

BREWSTER, S.A. (1994). *Providing a structured method for integrating non-speech audio into human-computer interfaces*. Ph.D. thesis, University of York, UK.

BREWSTER, S.A. (1997). Navigating telephone-based interfaces with earcons. In *Proceedings of HCI on People and Computers XII*, 39–56, Springer-Verlag, London, UK.

BREWSTER, S.A. (1998). Using nonspeech sounds to provide navigation cues. *ACM Trans. Comput.-Hum. Interact.*, **5**, 224–259.

BREWSTER, S.A. (2005). Sonically-enhanced widgets: Comments on Brewster and Clarke, ICAD 1997. *ACM Trans. Appl. Percept.*, **2**, 462–466.

BREWSTER, S.A. & EDWARDS, P.C.W.A.D.N. (1995). Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of HCI'95*, Huddersfield, UK.

BREWSTER, S.A., WRIGHT, P.C. & EDWARDS, A.D.N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of the INTERACT and CHI Conference on Human Factors in Computing Systems*, 222–227, ACM, New York, NY, USA.

BREWSTER, S.A., RÄTY, V.P. & KORTEKANGAS, A. (1996). Earcons as a method of providing navigational cues in a menu hierarchy. In *Proceedings of HCI on People and Computers XI*, 169–183, Springer-Verlag, London, UK.

BROWN, A., STEVENS, R. & PETTIFER, S. (2006). Audio representation of graphs: a quick look. In *Proceedings of the International Conference on Auditory Display*, Department of Computer Science, Queen Mary University of London, UK.

BROWN, L.M., BREWSTER, S., RAMLOLL, R., BURTON, M. & RIEDEL, B. (2003). Design guidelines for audio representation of graphs and tables. In E. Brazil & B. Shinn-Cunningham, eds., *Proceedings of the International Conference on Auditory Display*, 284–287, Boston University Publications Production Department, Boston, USA.

BUTCHER, J. (1992). *Copy-editing: The Cambridge Handbook for Editors, Authors and Publishers*, chap. 7: Preliminary pages, 177–179. Cambridge University Press, 3rd edn.

CAIRNS, P. & COX, A.L. (2008). *Research Methods for Human-Computer Interaction*. Cambridge University Press, New York, NY, USA.

CHANDRASHEKAR, S., STOCKMAN, T., FELS, D. & BENEDYK, R. (2006). Using think aloud protocol with blind users: a case for inclusive usability evaluation methods. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*, 251–252, ACM, New York, NY, USA.

CHILDS, E. (2005). Auditory graphs of real-time data. In E. Brazil, ed., *Proceedings of the International Conference on Auditory Display*, 402–405, Department of Computer Science and Information Systems, University of Limerick, Ireland.

CHIMERA, R. & SHNEIDERMAN, B. (1994). An exploratory evaluation of three interfaces for browsing large hierarchical tables of contents. *ACM Trans. Inf. Syst.*, **12**, 383–406.

COCKBURN, A., KARLSON, A. & BEDERSON, B.B. (2008). A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv.*, **41**, 1–31.

COHEN, R.F., MEACHAM, A. & SKAFF, J. (2006). Teaching graphs to visually impaired students using an active auditory interface. In *Proceedings of the SIGCSE Technical Symposium on Computer Science Education*, 279–282, ACM, New York, NY, USA.

COLLINS, K., KAPRALOS, B., HOGUE, A. & KANEV, K. (2010). An exploration of distributed mobile audio and games. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, 253–254, ACM, New York, NY, USA.

COLORADO STATE UNIVERSITY (2011). Writing guide: Writing abstracts. Retrieved on 28 January 2011, from http://writing.colostate.edu/ guides/documents/abstract/.

COOK, P.R., ed. (1999). *Music, Cognition and Computerized Sound: An Introduction to Psychoacoustics*. MIT Press, Cambridge, Massachussetts, United States.

COWAN, B. & KAPRALOS, B. (2008). Spatial sound for video games and virtual environments utilizing real-time gpu-based convolution. In *Proceedings of the Conference on Future Play: Research, Play, Share*, 166–172, ACM, New York, NY, USA.

CREASE, M. & LAU, T.Y. (2004). Evaluating the use of sound as a navigational aid on a mobile device - discussion. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services, Mobile HCI and Sound Workshop*, University of Strathclyde, Glasgow, Scotland.

DAVIES, T.C., BURNS, C.M. & PINDER, S.D. (2006). Using ecological interface design to develop an auditory interface for visually impaired travellers. In *Proceedings of the Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, 309–312, ACM, New York, NY, USA.

DAVIES, T.C., BURNS, C.M. & PINDER, S.D. (2007). Mobility interfaces for the visually impaired: what's missing? In *Proceedings of the ACM SIGCHI New Zealand chapter's International Conference on Computer-Human Interaction: Design Centered HCI*, 41–47, ACM, New York, NY, USA.

DESIGN COUNCIL (2011). Oxo good grips: design that everyone can use. Retrieved on 14 May 2011, from http://www.designcouncil.org.uk/case-studies/oxo-good-grips/.

DOMBOIS, F. (2001). Using audification in planetary seismology. In J. Hiipakka, N. Zacharov & T. Takala, eds., *Proceedings of the International Conference on Auditory Display*, 227–230, Laboratory of Acoustics and Audio Signal Processing and the Telecommunications Software and Multimedia Laboratory, Helsinki University of Technology, Espoo, Finland.

DONKER, H., KLANTE, P. & GORNY, P. (2002). The design of auditory user interfaces for blind users. In *Proceedings of the Nordic conference on Human-Computer Interaction*, 149–156, ACM, New York, NY, USA.

DOUSH, I.A. (2009). *Non Visual Navigation of Excel Spreadsheets*. Ph.D. thesis, New Mexico State University, USA.

EARL, J.D., CRISTA L.; LEVENTHAL (1999). A survey of windows screen reader users: Recent improvements in accessibility. *Journal of Visual Impairment & Blindness*, **93**.

EBERTS, R.E. (1994). *User interface design*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

EDWARDS, A.D.N. (1988). The design of auditory interfaces for visually disabled users. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 83–88, ACM Press.

EDWARDS, A.D.N., CHALLIS, B.P., HANKINSON, J.C.K. & PIRIE, F.L. (2000). Development of a standard test of musical ability for participants in auditory interface testing. In P.R. Cook, ed., *Proceedings of the International Conference on Auditory Display*.

EDWARDS, W.K. & MYNATT, E.D. (1994). An architecture for transforming graphical interfaces. In *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*, 39–47, ACM, New York, NY, USA.

EDWARDS, W.K., MYNATT, E.D. & RODRIGUEZ, T. (1993). *The Mercator Project: A Nonvisual Interface to the X Window System*. O'Reilly & Associates, Inc.

ESLAMBOLCHILAR, P. & MURRAY-SMITH, R. (2008). Interact, excite, and feel. In *Proceedings of the International Conference on Tangible and Embedded Interaction*, 131–138, ACM, New York, NY, USA.

FINLAYSON, J.L. (2006). Providing a tailored overview of program source code. In *CHI '06 extended abstracts on Human factors in computing systems*, 1743–1746, ACM Press, New York, NY, USA.

FINLAYSON, J.L. & MELLISH, C. (2005). The "audioview" – providing a glance at java source code. In *Proceedings of the International Conference on Auditory Display*, Department of Computer Science and Information Systems, University of Limerick, Ireland.

FLOWERS, J.H. (2005). Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions. In E. Brazil, ed., *Proceedings of the International Conference on Auditory Display*, 406–409, Department of Computer Science and Information Systems, University of Limerick, Ireland.

FOX, J., CARLILE, J. & BERGER, J. (2005). Sonimime: Sonification of fine motor skills. In E. Brazil, ed., *Proceedings of the International Conference on Auditory Display*, 316–318, Department of Computer Science and Information Systems, University of Limerick, Ireland.

FRAUENBERGER, C., DE CAMPO, A. & ECKEL, G. (2007). Analysing time series data. In G.P. Scavone, ed., *Proceedings of the International Conference on Auditory Display*, 504–508, Schulich School of Music, McGill University, Montreal, Canada.

FREEMAN, E. & GELERNTER, D. (1996). Lifestreams: a storage model for personal data. *SIGMOD Rec.*, **25**, 80–86.

FRIBERG, J. & GÄRDENFORS, D. (2004). Audio games: new perspectives on game audio. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, 148–154, ACM, New York, NY, USA.

GARZONIS, S., JONES, S., JAY, T. & O'NEILL, E. (2009). Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference. In *Proceedings of the International Conference on Human Factors in Computing Systems*, 1513–1522, ACM, New York, NY, USA.

GAVER, W.W. (1988). *Everyday listening and auditory icons*. Ph.D. thesis, University of California, San Diego.

GAVER, W.W. (1991). Sound support for collaboration. In *Proceedings of the European Conference on Computer-Supported Cooperative Work*, 293–308, Kluwer Academic Publishers, Norwell, MA, USA.

GAVER, W.W. (1993a). Synthesizing auditory icons. In *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*, 228–235, ACM, New York, NY, USA.

GAVER, W.W. (1993b). What in the world do we hear? an ecological approach to auditory event perception. *Ecological Psychology*, **5**, 1–29.

GAVER, W.W., SMITH, R.B. & O'SHEA, T. (1991). Effective sounds in complex systems: the arkola simulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology*, 85–90, ACM, New York, NY, USA.

GOOD, P. (1997). *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Springer, New York, NY, USA.

GUTWIN, C., SCHNEIDER, O., XIAO, R. & BREWSTER, S. (2011). Chalk sounds: the effects of dynamic synthesized audio on workspace awareness in distributed groupware. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 85–94, ACM, New York, NY, USA.

HARRAR, L. & STOCKMAN, T. (2007). Designing auditory graph overviews: An examination of discrete vs. continuous sound and the influence of presentation speed. In G.P. Scavone, ed., *Proceedings of the International Conference on Auditory Display*, 299–305, Schulich School of Music, McGill University, Montreal, Canada.

HERMANN, T. (2002). *Sonification for Exploratory Data Analysis*. Ph.D. thesis, Bielefeld University, Bielefeld, Germany.

HERMANN, T. (2008). Taxonomy and definitions for sonification and auditory display. In *Proceedings of the International Conference on Auditory Display*.

HERMANN, T. & RITTER, H. (1999). Listen to your data: Model-based sonification for data analysis. In G.E. Lasker, ed., *Advances in Intelligent Computing and Multimedia Systems*, 189–194, Int. Inst. for Advanced Studies in System research and cybernetics, Baden-Baden, Germany.

HERMANN, T. & RITTER, H. (2004). Sound and meaning in auditory data display. *Proceedings of the IEEE (Special Issue on Engineering and Music – Supervisory Control and Auditory Communication)*, **92**, 730–741.

HERMANN, T. & RITTER, H. (2005). Model-based sonification revisited—authors' comments on Hermann and Ritter, ICAD 2002. *ACM Trans. Appl. Percept.*, **2**, 559–563.

HERMANN, T., BAIER, G., STEPHANI, U. & RITTER, H. (2006). Vocal sonification of pathologic eeg features. In T. Stockman, L.V. Nickerson, C. Frauenberger, A.D.N. Edwards & D. Brock, eds., *Proceedings of the International Conference on Auditory Display*, 158–163, Department of Computer Science, Queen Mary University of London, UK.

HERMANN, T., BUNTE, K. & RITTER, H. (2007). Relevance-based interactive optimization of sonification. In G.P. Scavone, ed., *Proceedings of the International Conference on Auditory Display*, 461–467, Schulich School of Music, McGill University, Montreal, Canada.

HERMANN, T., BAIER, G., STEPHANI, U. & RITTER, H. (2008). Kernel regression mapping for vocal eeg sonification. In *Proceedings of the International Conference on Auditory Display*, Paris, France.

HEUTEN, W., HENZE, N. & BOLL, S. (2007). Interactive exploration of city maps with auditory torches. In *CHI Extended Abstracts on Human Factors in Computing Systems*, 1959–1964, ACM, New York, NY, USA.

HICKS, B.J., DONG, A., R.PALMER & MCALPINE, H.C. (2008). Organizing and managing personal electronic files: A mechanical engineer's perspective. *ACM Trans. Inf. Syst.*, **26**, 1–40.

HOCHHEISER, H. & LAZAR, J. (2010). Revisiting breadth vs. depth in menu structures for blind users of screen readers. *Interact. Comput.*, **22**, 389–398.

HORNBÆK, K. & FRØKJÆR, E. (2001). Reading of electronic documents: the usability of linear, fish-eye, and overview+detail interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 293–300, ACM, New York, NY, USA.

HORNBÆK, K., BEDERSON, B.B. & PLAISANT, C. (2002). Navigation patterns and usability of zoomable user interfaces with and without an overview. *ACM Trans. Comput.-Hum. Interact.*, **9**, 362–389.

HORNBÆK, K., BEDERSON, B.B. & PLAISANT, C. (2003). Navigation patterns & usability of zoomable user interfaces: with and without an overview. *interactions*, **10**, 11–12.

INTERNATIONAL COMMUNITY FOR AUDITORY DISPLAY (2011). Glossary. Retrieved on 28 January 2011, from http://icad.org/node/392.

JAGDISH, D., SAWHNEY, R., GUPTA, M. & NANGIA, S. (2008). Sonic grid: an auditory interface for the visually impaired to navigate gui-based environments. In *Proceedings of the International Conference on Intelligent User Interfaces*, 337–340, ACM, New York, NY, USA.

JONES, K.S. (2007). Automatic summarising: The state of the art. *Information Processing & Management*, **43**, 1449 – 1481.

JONES, M., JONES, S., BRADLEY, G., WARREN, N., BAINBRIDGE, D. & HOLMES, G. (2008). Ontrack: Dynamically adapting music playback to support navigation. *Personal Ubiquitous Comput.*, **12**, 513–525.

JOSEPH, A.J. & LODHA, S.K. (2002). Musart: Musical audio transfer function real-time toolkit. In R. Nakatsu & H. Kawahara, eds., *Proceedings of the International Conference on Auditory Display*, Advanced Telecommunications Research Institute (ATR), Kyoto, Japan.

KEIM, D.A. (1997). Visual techniques for exploring databases. Invited Tutorial, Int. Conference on Knowledge Discovery in Databases (KDD'97).

KILDAL, J. (2008). *Developing an Interactive Overview for Non-Visual Exploration of Tabular Numerical Information*. Ph.D. thesis, Scotland, UK.

KILDAL, J. & BREWSTER, S.A. (2005). Explore the matrix: browsing numerical data tables using sound. In *International Conference on Auditory Display*.

KILDAL, J. & BREWSTER, S.A. (2006a). Exploratory strategies and procedures to obtain non-visual overviews using tablevis. In *Intl Conf. Disability, Virtual Reality & Assoc. Tech*, 285–294.

KILDAL, J. & BREWSTER, S.A. (2006b). Non-visual overviews of complex data sets. In *Extended Abstracts of CHI*, 947, ACM Press.

KILDAL, J. & BREWSTER, S.A. (2006c). Providing a size-independent overview of non-visual tables. In *Proceedings of the International Conference on Auditory Display*.

KOOPMAN, P. (2011). How to write an abstract. Retrieved on 28 January 2011, from http://www.ece.cmu.edu/~koopman/ essays/abstract.html.

KRAMER, G. (1994). *Auditory Display: Sonification, Audification and Auditory Interfaces*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

KUBOVY, M. & VAN VALKENBURG, D. (2001). Auditory and visual objects. *Cognition*, **80**, 97–126.

KULES, B., KUSTANOWITZ, J. & SHNEIDERMAN, B. (2006). Categorizing web search results into meaningful and stable categories using fast-feature techniques. In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, 210–219, ACM, New York, NY, USA.

KUSTANOWITZ, J. & SHNEIDERMAN, B. (2005). Meaningful presentations of photo libraries: rationale and applications of bi-level radial quantum layouts. In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, 188–196, ACM, New York, NY, USA.

KWASNIK, B.H. (1992). A descriptive study of the functional components of browsing. In *Engineering for Human-Computer Interaction*, 191–203.

LANGLOIS, S., LOISEAU, S., TARDIEU, J., CERA, A. & MISDARIIS, N. (2010). Evaluation de la sonifcation d'un système multimédia automobile. In *Conference Internationale Francophone sur l'Interaction Homme-Machine*, 173–180, ACM, New York, NY, USA.

LEVINE, S.J. (2010). Writing and presenting your thesis or dissertation. Retrieved on 8 February 2011, from http://www.learnerassociates.net/ dissthes/#26.

LILJEDAHL, M., PAPWORTH, N. & LINDBERG, S. (2007). Beowulf: an audio mostly game. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, 200–203, ACM, New York, NY, USA.

LOHSE, J. (1991). A cognitive model for the perception and understanding of graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 137–144, ACM, New York, NY, USA.

LUK, J., PASQUERO, J., LITTLE, S., MACLEAN, K., LEVESQUE, V. & HAYWARD, V. (2006). A role for haptics in mobile interaction: initial design using a handheld tactile display prototype. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 171–180, ACM, New York, NY, USA.

MACKENZIE, I.S. & ZHANG, S. (1997). The immediate usability of graffiti. In *Proceedings of Graphics Interface*, 129–137, Toronto, Canada.

MAGNUSSON, C., MOLINA, M., RASSMUS-GRÖHN, K. & SZYMCZAK, D. (2010). Pointing for non-visual orientation and navigation. In *Proceedings of the Nordic Conference on Human-Computer Interaction: Extending Boundaries*, 735–738, ACM, New York, NY, USA.

MARSTON, J.R., LOOMIS, J.M., KLATZKY, R.L., GOLLEDGE, R.G. & SMITH, E.L. (2006). Evaluation of spatial displays for navigation without sight. *ACM Trans. Appl. Percept.*, **3**, 110–124.

MCGOOKIN, D.K. (2004). *Understanding and Improving the Identification of Concurrently Presented Earcons*. Ph.D. thesis, University of Glasgow.

MCGOOKIN, D.K. & BREWSTER, S.A. (2004a). Space, the final frontearcon: the identification of concurrently presented earcons in a synthetic spatialised auditory environment. In *International Conference on Auditory Display*, Sydney, Australia.

MCGOOKIN, D.K. & BREWSTER, S.A. (2004b). Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Trans. Appl. Percept.*, **1**, 130–155.

METATLA, O., BRYAN-KINNS, N. & STOCKMAN, T. (2008). Constructing relational diagrams in audio: the multiple perspective hierarchical approach. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*, 97–104, ACM, New York, NY, USA.

MILCZYNSKI, M., HERMANN, T., BOVERMANN, T. & RITTER, H. (2006). A malleable device with applications to sonification-based data exploration. In T. Stockman, L.V. Nickerson, C. Frauenberger,

A.D.N. Edwards & D. Brock, eds., *Proceedings of the International Conference on Auditory Display*, 69–76, Department of Computer Science, Queen Mary University of London, UK.

MILLER, D., PARECKI, A. & DOUGLAS, S.A. (2007). Finger dance: a sound game for blind people. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*, 253–254, ACM, New York, NY, USA.

MILLER, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, **63**, 81–97.

MILNE, A.P., ANTLE, A.N. & RIECKE, B.E. (2011). Tangible and body-based interaction with auditory maps. In *Proceedings of the Annual Conference Extended Abstracts on Human Factors in Computing Systems*, 2329–2334, ACM, New York, NY, USA.

MMWR (1996). Blindness caused by diabetes. *Morbidity and Mortality Weekly Report*, **45(43)**, 937–941.

MORELLI, T., FOLEY, J., COLUMNA, L., LIEBERMAN, L. & FOLMER, E. (2010). Vi-tennis: a vibrotactile/audio exergame for players who are visually impaired. In *Proceedings of the International Conference on the Foundations of Digital Games*, 147–154, ACM, New York, NY, USA.

MORENO, R. & MAYER, R.E. (1999). *Lecture Notes in Computer Science: Visual Information and Information Systems*, chap. Visual Presentations in Multimedia Learning: Conditions that Overload Visual Working Memory, 793–800. Springer Verlag, Berlin, Germany.

MURPHY, E. (2007). *Designing auditory cues for a multimodal web interface: a semiotic approach*. Ph.D. thesis, School of Music and Sonic Arts, Queen's University Belfast.

MURPHY, E., BATES, E. & FITZPATRICK, D. (2010). Designing auditory cues to enhance spoken mathematics for visually impaired users. In *Proceedings of the International ACM SIGACCESS conference on Computers and Accessibility*, 75–82, ACM, New York, NY, USA.

MYNATT, E.D. (1995). *Transforming graphical interfaces into auditory interfaces*. Ph.D. thesis, Georgia Institute of Technology.

MYNATT, E.D. & EDWARDS, W.K. (1992). Mapping GUIs to auditory interfaces. In *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*, 61–70, ACM Press.

MYNATT, E.D. & WEBER, G. (1994). Nonvisual presentation of graphical user interfaces: Contrasting two approaches. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 166–172, ACM Press.

NAKADA, M. & LEE DELLON, A. (1989). Relationship between sensibility and ability to read braille in diabetics. *Microsurgery*, **10**, 138–141.

NATIONAL FEDERATION OF THE BLIND JERNIGAN INSTITUTE (2009). The braille literacy crisis in america: Facing the truth, reversing the trend, empowering the blind. *Braille Monitor*.

NEES, M.A. & WALKER, B.N. (2008). Encoding and representation of information in auditory graphs: Descriptive reports of listener strategies for understanding data. In *Proceedings of the International Conference on Auditory Display*, Paris, France, inproceedings.

NICKERSON, L.V. (2007). Strategies for designing auditory overviews. Tech. Rep. RR-07-09, Department of Computer Science, Queen Mary University of London, UK.

NICKERSON, L.V. & HERMANN, T. (2008). Interactive sonifcation of grid-based games. In *Proceedings of the Audio Mostly Conference*, 27–34, Piteå, Sweden.

NICKERSON, L.V. & STOCKMAN, T. (2005). Sonically exposing the desktop space to non-visual users: An experiment in overview information presentation. In *Proceedings of the International Conference on Auditory Display*, Limerick, Ireland.

NICKERSON, L.V., STOCKMAN, T. & THIEBAUT, J.B. (2007). Sonifying the london underground real-time-disruption map. In G.P. Scavone, ed., *Proceedings of the International Conference on Auditory Display*, 252–257, Schulich School of Music, McGill University, Montreal, Canada.

NORTH, C., SHNEIDERMAN, B. & PLAISANT, C. (1996). User controlled overviews of an image library: a case study of the visible human. In *Proceedings of the first ACM international conference on Digital Libraries*, 74–82, ACM Press, New York, NY, USA.

OKTAY, B. & FOLMER, E. (2010). Textsl: a screen reader accessible interface for second life. In *Proceedings of the International Cross Disciplinary Conference on Web Accessibility*, 21:1–21:2, ACM, New York, NY, USA.

PAL, J., PRADHAN, M., SHAH, M. & BABU, R. (2011). Assistive technology for vision-impairments: anagenda for the ictd community. In *Proceedings of the International Conference Companion on World Wide Web*, 513–522, ACM, New York, NY, USA.

PALLADINO, D.K. & WALKER, B.N. (2007). Learning rates for auditory menus enhanced with spearcons versus earcons. In *Proceedings of the International Conference on Auditory Display*, 274–279.

PALLADINO, D.K. & WALKER, B.N. (2008). Navigation efficiency of two dimensional auditory menus using spearcon enhancements. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, **52**, 1262–1266(5).

PAPWORTH, N. (2010). ispooks: an audio focused game design. In *Proceedings of the Audio Mostly Conference: A Conference on Interaction with Sound*, 11:1–11:8, ACM, New York, NY, USA.

PARENTE, P. (2005). Clique: A conversant, task-based audio display for GUI applications. In *Doctoral Consortium ASSETS'05*.

PARENTE, P. (2008). *Clique: perceptually based, task oriented auditory display for gui applications*. Ph.D. thesis, NC, USA, aAI3322224.

PARENTE, P. & BISHOP, G. (2003). Bats: The blind audio tactile mapping system. In *Proceedings of ACM South Eastern Conference*.

PARENTE, P. & BISHOP, G. (2009). Out from behind the curtain: learning from a human auditory display. In *Proceedings of the International Conference Extended Abstracts on Human Factors in Computing Systems*, 2575–2584, ACM, New York, NY, USA.

PARENTE, P. & CLIPPINGDALE, B. (2006). Linux screen reader: extensible assistive technology. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*, 261–262, ACM, New York, NY, USA.

PATERSON, N., NALIUKA, K., JENSEN, S.K., CARRIGY, T., HAAHR, M. & CONWAY, F. (2010). Design, implementation and evaluation of audio for a location aware augmented reality game. In *Proceedings of the International Conference on Fun and Games*, 149–156, ACM, New York, NY, USA.

PAULETTO, S. & HUNT, A. (2005). A comparision of audio & visual analysis of complex time-series data sets. In E. Brazil, ed., *Proceedings of the International Conference on Auditory Display*, 175–181, Department of Computer Science and Information Systems, University of Limerick, Ireland.

PETRIE, H., HAMILTON, F. & KING, N. (2004). Tension, what tension?: Website accessibility and visual design. In *Proceedings of the International Cross-Disciplinary Workshop on Web Accessibility*, 13–18, ACM, New York, NY, USA.

PETRUCCI, L.S., HARTH, E., ROTH, P., ASSIMACOPOLOUS, A. & PUN, T. (2000). Websound: A generic web sonification tool, and its application to an auditory web browser for blind and visually impaired users. In P.R. Cook, ed., *Proceedings of the International Conference on Auditory Display*, Georgia Institute of Technology, Atlanta, GA, USA.

PIELOT, M., HENZE, N., HEUTEN, W. & BOLL, S. (2007). Tangible user interface for the exploration of auditory city maps. In *Proceedings of the International Conference on Haptic and Audio Interaction Design*, 86–97, Springer-Verlag, Berlin, Heidelberg.

PITT, I.J. & EDWARDS, A.D.N. (1991). *People and Computers VI: Proceedings of HCI '91*, chap. Navigating the Interface by Sound for Blind Users, 373–383. Cambridge University Press.

PLAISANT, C. & SHNEIDERMAN, B. (1995). Organization overviews and role management: inspiration for future desktop environments. In *Conference Companion on Human Factors in Computing Systems*, 419–420, ACM, New York, NY, USA.

PLAISANT, C., MILASH, B., ROSE, A., WIDOFF, S. & SHNEIDERMAN, B. (1996). Lifelines: visualizing personal histories. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 221–227, ACM, New York, NY, USA.

PLAISANT, C., HELLER, D., LI, J., SHNEIDERMAN, B., MUSHLIN, R. & KARAT, J. (1998a). Visualizing medical records with lifelines. In *CHI conference summary on Human factors in computing systems*, 28–29, ACM, New York, NY, USA.

PLAISANT, C., MARCHIONINI, G. & KOMLODI, A. (1998b). Bringing treasures to the surface: previews and overviews in a prototype for the library of congress national digital library. In *CHI Conference summary on Human Factors in Computing Systems*, 187–188, ACM, New York, NY, USA.

PÉREZ-QUIÑONES, M.A., CAPRA, R.G. & SHAO, Z. (2003). The ears have it: a task by information structure taxonomy for voice access to web pages. In *Proc. IFIP Interact*, IOS Press.

RAMAN, T.V. (1996a). Emacspeaka speech interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground*, 66–71, ACM, New York, NY, USA.

RAMAN, T.V. (1996b). Emacspeakdirect speech access. In *Proceedings of the Annual ACM Conference on Assistive Technologies*, 32–36, ACM, New York, NY, USA.

RAMLOLL, R., YU, W., BREWSTER, S., RIEDEL, B., BURTON, M. & DIMIGEN, G. (2000). Constructing sonified haptic line graphs for the blind student: first steps. In *Proceedings of the International ACM Conference on Assistive Technologies*, 17–25, ACM, New York, NY, USA.

RÖBER, N. & MASUCH, M. (2004). Interacting with sound: An interaction paradigm for virtual auditory worlds. In *Proceedings of the International Conference on Auditory Display*, Sydney, Australia.

SÁNCHEZ, J., SAENZ, M. & GARRIDO, J.M. (2010a). Usability of a multimodal video game to improve navigation skills for blind children. *ACM Trans. Access. Comput.*, **3**, 7:1–7:29.

SÁNCHEZ, J., SÁENZ, M., PASCUAL-LEONE, A. & MERABET, L. (2010b). Enhancing navigation skills through audio gaming. In *Proceedings of the of the International Conference Extended Abstracts on Human Factors in Computing Systems*, 3991–3996, ACM, New York, NY, USA.

SATO, D., ZHU, S., KOBAYASHI, M., TAKAGI, H. & ASAKAWA, C. (2011). Sasayaki: augmented voice web browsing experience. In *Proceedings of the Annual Conference on Human Factors in Computing Systems*, 2769–2778, ACM, New York, NY, USA.

SAVIDIS, A., STEPHANIDIS, C., KORTE, A., CRISPIEN, K. & FELLBAUM, K. (1996). A generic direct-manipulation 3D-auditory environment for hierarchical navigation in non-visual interaction. In *Proceedings of the Annual ACM Conference on Assistive Technologies*, 117–123, ACM Press.

SAWHNEY, N. & SCHMANDT, C. (1999). Nomadic radio: scaleable and contextual notification for wearable audio messaging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI is the Limit*, 96–103, ACM, New York, NY, USA.

SAWHNEY, N. & SCHMANDT, C. (2000). Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Trans. Comput.-Hum. Interact.*, **7**, 353–383.

SENN, S. (2002). *Cross-over trials in clinical research (statistics in practice)*. Wiley, Chichester, UK.

SHNEIDERMAN, B. (1994). Dynamic queries for visual information seeking. *IEEE Software*, **11**, 70–77.

SHNEIDERMAN, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proc. IEEE Conference on Visual Languages*, 336–343, Boulder, CO, USA.

SMITH, D.A. & WALKER, B.N. (2002). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. In R. Nakatsu & H. Kawahara, eds., *Proceedings of the International Conference on Auditory Display*, Advanced Telecommunications Research Institute (ATR), Kyoto, Japan.

SMITH, M.A., SHNEIDERMAN, B., RODRIGUES, N.M.F.E.M., BARASH, V., DUNNE, C., CAPONE, T., PERER, A. & GLEAVE, E. (2009). Analyzing (social media) networks with NodeXL. In *Proceedings of the International Conference on Communities and Technologies*, 255–264, ACM, New York, NY, USA.

SONNEK, P. & RICE, S.V. (2010). Synthesizing the acoustic doppler effect in software. In *Proceedings of the Annual Southeast Regional Conference*, 71:1–71:5, ACM, New York, NY, USA.

SPILIOTOPOULOS, D., STAVROPOULOU, P. & KOUROUPETROGLOU, G. (2009). Acoustic rendering of data tables using earcons and prosody for document accessibility. In *Proceedings of the International Conference on Universal Access in Human-Computer Interaction. Part III: Applications and Services*, 587–596, Springer-Verlag, Berlin, Heidelberg.

STEVENS, R.D. (1996). *Principles for the design of auditory interfaces to present complex information to blind people*. Ph.D. thesis, University of York, UK.

STEVENS, R.D., BREWSTER, S., WRIGHT, P.C. & EDWARDS, A.D.N. (1994). Design and evaluation of an auditory glance at algebra for blind readers. 21–30, International Community for Auditory Display, International Community for Auditory Display, Santa Fe, NM, U.S.

STEVENS, R.D., WRIGHT, P.C., EDWARDS, A.D.N. & BREWSTER, S.A. (1996). An audio glance at syntactic structure based on spoken form. In *Proceedings of the 5th international conference on Computers helping people with special needs. Part II*, ICCHP '96, 627–635, R. Oldenbourg Verlag GmbH, Munich, Germany, Germany.

STOCKMAN, T. (2005). Interactive sonification of spreadsheets. In *Proceedings of the International Conference on Auditory Display*.

STOCKMAN, T. & METATLA, O. (2008). The influence of screen-readers on web cognition. In *Proceedings of Accessible Design in the Digital World*, York, UK.

STRACHAN, S., ESLAMBOLCHILAR, P., MURRAY-SMITH, R., HUGHES, S. & O'MODHRAIN, S. (2005). Gpstunes: controlling navigation via audio feedback. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices & Services*, 275–278, ACM Press, New York, NY, USA.

STRAIN, P., SHAIKH, A.D. & BOARDMAN, R. (2007). Thinking but not seeing: think-aloud for non-sighted users. In *CHI Extended Abstracts on Human Factors in Computing Systems*, 1851–1856, ACM, New York, NY, USA.

THEOFANOS, M.F. & REDISH, J.G. (2003). Bridging the gap: between accessibility and usability. *interactions*, **10**, 36–51.

TREWIN, S., LAFF, M., HANSON, V. & CAVENDER, A. (2009). Exploring visual and motor accessibility in navigating a virtual world. *ACM Trans. Access. Comput.*, **2**, 11:1–11:35.

TULVING, E. & PEARLSTONE, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, **5**, 381 – 391.

TVERSKY, B. (2000). *Cognitive Mapping - Past, present and future*, chap. Levels and structure of spatial knowledge, 24–43. Routledge.

UNIVERSITY OF NORTH CAROLINA (2011). Abstracts. Retrieved on 28 January 2011, from http://www.unc.edu/depts/wcweb/ handouts/abstracts.html.

USABILITY FIRST (2011). Usability first - usability glossary - focus+context | usability first. Retrieved on 28 January 2011, from http://www.usabilityfirst.com/ glossary/focuscontext/.

VAN DER VEER, G.C. & DEL CARMEN PUERTA MELGUIZO, M. (2003). In J.A. Jacko & A. Sears, eds., *The Human-Computer Interaction Handbook*, chap. Mental models, 52–80, L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

VICKERS, P. & ALTY, J.L. (2005). Musical program auralization: Empirical studies. *ACM Trans. Appl. Percept.*, **2**, 477–489.

WALKER, B.N. & LANE, D.M. (2001). Sonification mappings database on the web. In J. Hiipakka, N. Zacharov & T. Takala, eds., *Proceedings of the International Conference on Auditory Display*, 281, Laboratory of Acoustics and Audio Signal Processing and the Telecommunications Software and Multimedia Laboratory, Helsinki University of Technology, Espoo, Finland.

WALKER, B.N. & LINDSAY, J. (2003). Effect of beacon sounds on navigation performance in a virtual reality environment. In *International Conference on Auditory Display*, Boston, Massachussetts, United States.

WALKER, B.N., NANCE, A. & LINDSAY, J. (2006). Spearcons: speech-based earcons improve navigation performance in auditory menus. In T. Stockman, L.V. Nickerson, C. Frauenberger, A.D.N. Edwards & D. Brock, eds., *Proceedings of the International Conference on Auditory Display*.

WEBER, H., ed. (1984). *Webster's II New Riverside Dictionary*. Berkley Books.

WINBERG, F. & HELLSTROM, S.O. (2003). Designing accessible auditory drag and drop. In *Proceedings of the Conference on Universal Usability*, 152–153, ACM, New York, NY, USA.

WORRALL, D. (2009). The use of sonic articulation in identifying correlation in capital market trading data. In *Proceedings of the International Conference on Auditory Display*, Re:New Digital Arts Forum, Copenhagen, Denmark.

YEO, W.S. & BERGER, J. (2005). Framework for designing image sonification methods. In E. Brazil, ed., *Proceedings of the International Conference on Auditory Display*, 323–327, Department of Computer Science and Information Systems, University of Limerick, Ireland.

YU, W. & BREWSTER, S. (2002). Comparing two haptic interfaces for multimodal graph rendering. In *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 3–, IEEE Computer Society, Washington, DC, USA.

YU, W., MCALLISTER, G., STRAIN, P., KUBER, R. & MURPHY, E. (2005). Improving web accessibility using content-aware plug-ins. In *CHI Extended Abstracts on Human Factors in Computing Systems*, 1893–1896, ACM, New York, NY, USA.

YU, Y. & LIU, Z. (2010a). Improving the performance and usability for visual menu interface on mobile computers. In *Proceedings of the International Conference on Advanced Visual Interfaces*, 369–372, ACM, New York, NY, USA.

YU, Y. & LIU, Z. (2010b). A user study of visual versus sonically-enhanced interfaces for use while walking. In *Proceedings of the International Conference on Multimedia*, 687–690, ACM, New York, NY, USA.

YUAN, B. & FOLMER, E. (2008). Blind hero: enabling guitar hero for the visually impaired. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*, 169–176, ACM, New York, NY, USA.

ZHAO, H. (2006). *Interactive sonification of abstract data: framework, design space, evaluation, and user tool*. Ph.D. thesis, University of Maryland at College Park, College Park, MD, USA.

ZHAO, H., PLAISANT, C., SHNEIDERMAN, B. & DURAISWAMI, R. (2004). Sonification of geo-referenced data for auditory information seeking: Design principle and pilot study. In *Proceedings of the International Conference on Auditory Display*, Sydney, Australia.

ZHAO, H., SHNEIDERMAN, B. & PLAISANT, C. (2006). iSonic: Interactive data sonification for blind users. Retrieved on 28 January 2011, from http://www.cs.umd.edu/hcil/audiomap.

ZHAO, H., PLAISANT, C., SHNEIDERMAN, B. & LAZAR, J. (2008). Data sonification for users with visual impairment: A case study with georeferenced data. *ACM Trans. Comput.-Hum. Interact.*, **15**, 4:1–4:28.