UNIVERSITY
of
GLASGOW

Department of
Computing Science

The design and evaluation of non-speech sounds to support navigation in restricted display devices

Grégory Leplâtre

Submitted for the degree of Doctor of Philosophy
to
The University of Glasgow,
November, 2002.

Abstract

This thesis proposes a principled approach to the design of non-speech audio to support interaction in limited display devices. Prior to this research, no principles had been devised to systematically design sounds that support navigation in large mobile phone menus, and consequently no sonification had been successfully designed to support navigation in such interfaces.

The design principles in this thesis are supported by implementation guidelines and were applied in two case studies. The first application involves the development of a sonification for a mobile phone prototype in which the sounds can only be generated by the telephone buzzer, which constitutes a serious constraint on the design. The second application involves the sonification of a similar menu with no technical constraint on the design of sounds.

The sonification designed for the second application was evaluated and demonstrated that the sounds improved the performance of navigational tasks in the menu in many respects. These results confirm the soundness of the design and hence of the design framework. They also prove that non-speech sounds can improve the usability of large menus in restricted display devices.

This thesis also tackles the practical issues involved in designing auditory interfaces. A formalism that facilitates the design and integration of sounds in a user interface is presented. AIDE, an Auditory Interface Design Environment that partly meets the requirements of this formalism provides a concrete illustration of this reflection.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank all the people in the Department fo Computing Science at the University of Glasgow who have assisted in the research that led to this thesis. In particular, my supervisor, Stephen Brewster for his insightful guidance throughout the course of this research.

I would also like to thank Seppo, Outti, Yarkko and their colleagues from Nokia, whose interest in this research came as a great motivational factor when I most needed it.

Thanks to my friends, in the department or elsewhere, for having made these years in Glasgow a great experience.

Thanks to my family for their unconditional support and love.

And thanks to my Nur ...

This work was funded by EPSRC grant GR/L66373.

à mes parents…

Declaration

The experiment described in Chapter 4 has been published at ICAD'98 [LB98]. The evaluation of the mobile-telephone prototype sonification presented in Chapter 6 was published at ICAD'2001 [HLML01]. The experiment described in Chapter 7 was published at ICAD'2000 [LB00]. This thesis only exploits those parts of collaborative papers that are directly attributable to the author.

Chapter 1

Introduction

1.1 Context

With the explosion of mobile telecommunications, telephones are truly ubiquitous. Mobile and fixed telephones not only allow people to communicate, they are also a common means of accessing various forms of information and services, for example telephone banking and checking email on a WAP telephone. Nevertheless, telephone-banking, WAP[1] services or navigating in a mobile telephone menu is the cause of many users' frustration. As one journalist noted in 1990, "(interactive voice response systems) deserve A+ for irritation" [How90]. The aim of this thesis is to provide a solution to this problem.

1.1.1 Interaction with Restricted Display Devices

Even though there are significant differences between a traditional telephone and a mobile telephone, it is arguable that these devices suffer from the same problems: allowing users to carry out complex tasks with a limited interaction bandwidth. No matter how good the design of a mobile telephone menu or a WAP service is, interaction is still constrained by the inherent limitations of these devices. These include a small screen and poor input techniques.

The most common method for the functionality of a service or a device to be laid out is through a hierarchical menu. Due to the large number of options available, these menus are

[1]WAP (Wireless Application Protocol) is the de facto worldwide standard for providing Internet communications and advanced telephony services on digital mobile telephones, pagers, personal digital assistants and other wireless terminals (see http://www.wapforum.org)

very often large and have complex structures. Previous research has shown that, because of the limitations of these devices, navigation in telephone-based menus is the main cause of interaction problems [Ros85, WKK95].

The recent improvements in the display qualities of mobile devices suggest that interaction with such devices can be improved through graphics. However, the size of the screen is inevitably limited on these devices, because they must be small enough to fit in a pocket, which suggests that improvements related to graphical display will remain limited. In addition, there will always be some categories of devices that do not contain any visual output. For example, most traditional telephones do not, and the trends in the development of the next generation of mobile devices demonstrate that the multiplication of device categories will lead to the development of small, screenless, devices, for instance pocket MP3[2] players.

1.1.2 Supporting Interaction with Non-Speech Sounds

While graphical displays may fail to present information to users in a usable fashion, the auditory modality remains either unused or solely used for the purposes of speech output. Yet non-speech sounds have proved successful at conveying information to users in various forms [AVR97, Bre94, Gav97].

Previous research has tackled the issue of designing sounds for hierarchical menus, but it is arguable that these designs have only partly addressed the issue of navigaton. Brewster and colleagues have investigated the use of non-speech sounds to communicate information regarding the structure of small hierarchical menus [BRK96, Bre97]. These studies demonstrate that the structure of a small hierarchical menu can be mapped to a set of hierarchically designed non-speech sounds and that users can locate the sounds in the hierarchy after a period of familiarisation with the menu. However, more work is required before sounds can be implemented in real-life menus. The main issues that need to be addressed are:

- Do non-speech sounds improve the performance of practical navigational tasks?
- Is it possible to design non-speech sounds for a real-life menu hierarchy, such as a mobile telephone menus containing hundreds of nodes?

[2] MP3 files are compressed audio files.

1.2 Subject and Research Approach

This thesis proposes a new approach to supporting navigation with non-speech sounds, based on the nature of navigating in a hierachical menu. The premise of this approach is a thorough understanding of navigation in telephone-based menus. The core of the framework proposed involves a set of design principles and implementation recommendations, that will guide designers through the process of devising a *sonification*[3].

The main application of this framework addresses the two questions already raised: a set of non-speech sounds has been devised for a large mobile telephone menu (The NOKIA *6110* menu) from new design principles and guidelines. In addition, the evaluation of this sonification involved actual navigation tasks. The effectiveness of the design was measured by comparing the performance of the tasks by a group of users of a standard mobile telephone menu with the performance of the same tasks by a group of users of the sonified menu.

An important issue related to sound design for user interfaces is that of creating prototypes of the auditory interface. Tools developed to make music offer a rich environment to create audio materials, but it is difficult, if not impossible, to link them to a user interface prototype. On the other hand, the prototyping tools available on the user interface side (Macromedia Director for example) do not offer sufficient audio resources to carry out complex sonifications. Therefore, the question of complex auditory interface prototyping will also be addressed in this thesis through the development of a novel design tool.

1.3 Structure of the Thesis

The next two chapters of this thesis will be devoted to the presentation of the background research issues and knowledge required to follow the research path proposed so far in this introduction. These research issues can be divided into two categories: research concerning audio and telephone-based interfaces (these interfaces are the focus of the work presented in this thesis) and research on the perception of sound. In particular, these sections will review the most important questions that concern this research:

- What is navigation?
- How to support navigation with non-speech audio?

[3] See Section 3.6.1 for a hierarchical menu.

After this review of relevant research, a preliminary experimental study will make the junction between previous research by Brewster and colleagues (see [BRK96]) and the new approach described in this thesis. This study will investigate new design parameters to represent hierarchical menus in sound (See Chapter 4).

Chapter 5 will then present the theoretical part of the research: a set of design principles will be devised from an analysis of the tasks involved in menu navigation. These principles will be supported by relevant implementation recommendations.

The first application of these principles and implementation guidelines will be presented in Chapter 6. This section will describe the sonification of a mobile telephone prototype based on the NOKIA *8210*. The motivations for this study are multiple:

- It offers an opportunity to apply the principles proposed in Chapter 5 to a real-life protoype.
- The technical requirements of the mobile telephone prototype impose many constraints on the type of sounds allowed in the sonification (the sounds could only be generated by the mobile telephone buzzer). These constraints were used as a healthy challenge to the design principles.

The main application of the design framework devised in Chapter 5 will be presented in Chapter 7. The application is similar to that described above in that it concerns a large mobile telephone menu, but this time the sonification ran on a desktop simulation of the mobile telephone menu. However, no constraints were imposed on the types of sounds used for this sonification. The audio quality of the new generation of mobile devices improving constantly, this application constitutes the most interest for the long term. In addition, a detailed evaluation of the sonification will be presented. The strength of this evaluation relies on the fact that it involves realistic navigation tasks. The perfomance of these tasks will be measured in both audio and non-audio conditions. These results will either confirm or deny the soundness of the design and consequently the soundness of the design framework proposed in Chapter 5.

This thesis will conclude with a reflection on the practical sound design issues faced during the course of this research. The large scale design carried out in Chapter 7 involved over a hundred different sounds, showed the limits of using tools such as sequencers to design auditory interfaces. Chapter 8 will propose a formalism that facilitates the design and integration of sounds in a user interface. The theoretical approach to the problem will be completed by the presentation of AIDE, an Auditory Interface Design Environment that partly meets the requirements of the formalism.

1.4 CD provided with the thesis

This thesis addresses a number of issues that require multimedia illustrations. The CD provided with this thesis contains all the multimedia materials referenced in this thesis: audio samples, telephone menu simulations and software. Appendix D provides directions on how to access and use these.

Chapter 2

Telephone-based interfaces

2.1 Introduction

This present chapter reviews the interaction issues related to telephone-based interfaces (TBIs). Firstly, it gives an overview of the various types of interfaces that fall under the category of telephone-based interfaces. After presenting the two main ways that people interact with a telephone (menu-based interaction versus command-based interaction), this chapter will focus on the chief aspect of interaction investigated in this thesis: navigation. The concept of navigation will be reviewed, firstly in general and then with a focus on telephone menus.

2.1.1 Use of Telephone-Based Interfaces

The telephone is a widespread means to access a large variety of services. Although most people admit that these systems are useful, few people like to use telephone-based interactive services. A decade ago, a journalist clearly reflected people's annoyance when they use telephone-based interactive systems. His claim is still valid today:

> "The explosion of telecommunication devices has turned the telephone into anything but a user-friendly instrument: As a matter of fact these days, placing a telephone call, once thought as a routine, is more is more like taking a trip through a mine field. Topping my personal list of annoying interventions is the increasingly popular attendant hookup (*i.e.*, interactive voice response systems), which certainly deserves a A+ for irritation" [How90].

The term TBI refers to a broad range of interaction situations, from the access to a remote service from a fixed telephone, to the performance of a local operation on a mobile device. As pointed out by Karis [KZ89], mobile telephony is fairly distinct from traditional telephony. Indeed mobile telephony and traditional telephony exhibit very different interaction issues: the latter almost exclusively involves non-visual interaction while the former involves navigation in graphical menus as well.

With the rapid expansion of mobile telephony, the principal human factors studies tend to focus on the use of telephone handsets to access remote information. For instance, Goose *et al.* have studied the possibility of using telephones to browse hypertext [GWM98]; Marx and Schmandt have developed a conversational system that allowed users to send and receive emails through the telephone [MS96b, MS96a]. Miller and Elias have investigated the use of menu-based interface to access a remote computer with a telephone [ME91], whereas Schmandt developed a similar application, but based on a human/machine conversation model [Sch93].

In what follows, TBIs will refer to the telephone interface in the most general sense. It may involve exclusively auditory interaction, such as interacting vocally with a speech-recognition system, it may also involve navigation in a graphical mobile telephone menu. The distinction between these interaction styles will be made when necessary.

2.1.2 Interaction Problems with TBIs

Schumacher *et al.* explain that interaction problems with TBIs are a consequence of the increased availability of services with a lack of design standardisation [SHS95]. Human factors studies generally agree on the fact that the increased availability of more sophisticated systems has led to a number of divergent interface styles [Mag96, Sch92, SHS95]. Nevertheless, Maguire states that complete convergence around a single set of standards for telephone interfaces is neither desirable nor achievable [Mag96]. Indeed, given the breadth of interface types available, it is preferable to focus on families of interfaces in order to address interaction issues.

Interaction problems with TBIs are not only the result of the increased availability of services or functions on telephones. In addition, this expansion of functionality occurs on a device whose interaction bandwidth[1] is extremely narrow. Schumacher *et al.* highlight the main physical restrictions of the device [SHS95]:

[1] Interaction bandwidth typically refers to richness of the input/output available with a given interface.

- Auditory interaction only.

- Telephone-based interfaces push the limit on users' working memory. Because all output is auditory, it is presented serially through the telephone hand-set. The serial presentation of information places heavy demand on working memory.

- Limited input device. Telephone interface designers must assume that people use telephones that have only 12 touch-tone keys.

Schumacher's description clearly applies to 'traditional' telephones only. Most mobile telephones now allow visual interaction in certain situations. However, the distinction does not come from the device itself but from the application or service running on the device. The following categorisation is useful to identify the various interaction situations that fall under interacting with TBIs:

Category 1 Speech-based menu. This concerns the widespread telephone-based services such as telephone banking. The output is invariably speech and the input can be either the telephone keypad, or speech, or a combination of both.

Category 2 Hierarchical menu. This category is related to navigation in mobile telephone menus. The output can be text-based only or involve graphics and varies greatly depending on the device (for example, the Ericsson T20 mobile telephone menu is exclusively text-based, while the Sony J5 contains some graphical 3D effects in its main menu). The type of input required also varies greatly depending on the device. The keypad is used to enter numbers and letters; navigation is carried out with arrow keys and sometimes a 'roller' which makes browsing long lists of menus faster. Most of today's telephones have a text-completion functionality where a syllable or word morpheme can be written automatically by keying in its first letter. In this way, writing text using the keypad is speeded up.

Category 3 Hypermedia information structure. This category, similar to the previous one, is typical of WAP telephones. The main difference between the current category and the previous one lies in the fact that WAP content is not necessarily hierarchical. In addition, WAP does not support features such as audio output whereas some mobile telephone menus do. The input for this category is usually also through the keypad and roller[2].

[2] The text completion functionality however is rarely included in today's WAP browsers

Category 4 Hyper-and-multi-media information structure. This last category is related to the so-called "third generation" of mobile devices. This type of device offers the possibility of accessing similar content to that available on today's networked computers but, obviously, using a much more compact type of device. Differences from WAP include good quality of graphics and auditory output. The input for this generation of device is very uncertain but will necessarily be constrained by the size of the devices.

The only usability studies available today are essentially related to the first three categories defined above. Due to the importance of traditional telephony, a large number of studies relate to the first two categories shown above [SH93, Sch92, SHS95]. The main issue identified by these studies is that telephone features have to share space within the limited interaction bandwidth of such devices. According to Griffeth, the interaction between telephone features is responsible for making a simple interface (the telephone) a confusing interface [Gri96]. As a result, the interface does not meet the criteria defining well-designed objects. These criteria are: the user knows what to do and the user can tell what is happening [Nor90].

As Roberts *et al.* suggest, the problems experienced by users while using a telephone-based interface are not the consequence of a bad design of the system, but result from the nature of the technology itself [RE89]. A typical example of one drawback subsequent to the proliferation of telephone functionality is the use of the same feedback for different purposes. For instance, there are three different reasons why telephones can ring: Someone (or some computer) is calling; the user has hung up, but the call is on hold; or a monitored line is now available and a call can be made to it. Griffeth concludes by pointing out that there is a real need for human factors research in this area [Gri96]:

> "The challenge for human factors research is to determine how one resolves interactions in a manner natural to users of a communication device. Even recognizing that only user studies can make the ultimate determination of design quality, a collection of principles that one can follow to develop possible solutions would represent a major advance".

Deffner and Melder suggest that ultimately, TBIs designers should aim at meeting the following criteria while creating an interface [DM90]:

- Letting the user know where he/she is within the system.

- Letting the user know what functionality is available at any given moment.
- Giving the user quick access to desired functionality.

2.1.3 Means of Interacting with TBIs

There are many different types of TBIs and many ways of classifying these interfaces. In the previous section a categorisation of TBIs based on interaction bandwidth was proposed. The present section presents the two main interaction styles available for TBIs in all of the categories already described: command-based interfaces and menu-based interfaces. The fact that these are dealt with in two distinct sections does not mean that they could not cohabit in the same system. On the contrary, the future of non-visual interfaces probably lies in the adequate combination of both.

Visual versus Non-Visual Interaction

When an interface contains some form of visual display, it is common practice in interface design to convey as much information as possible through the visual channel, even if an auditory channel is available. As today's interaction styles use visual output, there might be a case for ignoring the auditory channel. However, there are several reasons why the auditory channel should not be neglected:

- The auditory channel is the only available channel for a large category of interfaces (Category 1 in Section 2.1.2).
- When visual output is available, there is only so much information that can be displayed on a small screen. The audio channel can be used to convey additional information.
- Even though most mobile telephones and mobile devices have a screen, smaller devices with no screen will appear on the market in the future. Interacting with such devices would involve auditory output only.
- Mobile devices' screens are not usable when the device is at the user's ear.

The fundamental difference between visual and auditory output lies in the way information is displayed over time. Visual output can present information in a static form, which users can look at whenever they want. On the other hand, audio information is necessarily dynamic and users cannot easily control the flow of information presented to them. Therefore, effective presentation of auditory information over a short time interval is a crucial factor [ME97]. Compared to visual presentation of information, speech output is slow, serial and provides no short-term memory aids [HN89]. As Resnick and Virzi mention, good readers can read faster than they can listen [RV95]. The temporal factor is thus the critical one that distinguishes visual interfaces from non-visual ones. Studies have shown that this factor also prevents designers from applying menu design guidelines valid in a visual context, to an auditory interface [MLL86, PPS92]. Resnick and Virzi's view of the future of temporal interaction sums up the challenge that needs to be met:

> "The temporal presentation of audio creates interesting design challenges. Exploration of techniques that give users control over the time dimension is just beginning. The essence of that exploration is to break information chunks into even-smaller parts and to find natural ways for users to control which part will be presented next" [RV95].

Nevertheless, auditory display can have advantages over visual display which one could benefit from. For instance, a study by Savage showed that certain tasks performed using auditory help messages had a significantly lower error rate than did those tasks performed using visual instructions [SK91]. More information on the use of sounds in human-computer interaction will be provided in the next chapter

Finally, from a broader point of view, it should be noted that using menus to interact with a system in a non-visual environment, or having a conversation with a system may not be the only means of interaction. Is there an equivalent of the desktop metaphor for non-visual interfaces? Can new metaphors be envisaged that would make auditory interaction easier? There are no satisfying answers to these questions available in the relevant literature; it is arguable that it would be as important for the future of non-visual interfaces to investigate these issues as it is to focus on existing interaction methods. As these questions go beyond the scope of the present thesis, they are only raised for the reader's awareness.

Because TBIs cover a large number of devices and interaction categories, the interaction style review will be carried out from the least common denominator point of view. The focus will be on the most constrained type of interaction: the first interaction category (see Section 2.1.2). In the next two sections, unless stated otherwise, the term TBI will refer to this category of interaction.

2.2 Menu-Based Interfaces

Menu-based interfaces are currently the most common means of interaction with a TBI. The output of the interface is made from spoken prompts or text (interaction categories 1, 2 and 3), and the input is either automatic speech recognition (ASR) or touch-tone (TT).

Menu-based interface design can be divided into two parts: menu design and dialogue design. These two parts are not completely independent. Menu design refers to the determination of the structure of the menu *e.g.*, how the hierarchy should be built, how menu prompts should be organised within the hierarchy. Menu dialogue refers to the means by which users can navigate through the hierarchy *e.g.*, how the interaction with the system takes place.

2.2.1 Menu Design

According to Schwartz and Hardzinski, the overall structure should not involve more than three levels with more than four prompted choices at each level [SH93]. However, *Skip and scan* menus are an alternative to traditional menus, these allow the number of menu items available to be extended (interaction category 1 only) [RV92]. Skip and scan menus are two-button type menus that allow users to quickly skip options with one button when necessary, and make a selection with the other button. These should be used when more than the recommended four prompts are used at one level of a menu. For instance, to display a list of 15 films in a movie booking service, it is preferable to give the user the option of skipping irrelevant information and selecting the item desired.

Searching is a fundamental task that has been investigated to measure the usability of telephone menus. Martin *et al.* report from an experiment that the performance of a given search task is affected by the structure of the menu [MWW90]. In this study, a wide and shallow menu (eight items at each node on two levels) was compared with a deep and narrow menu (two items per node on six levels). Participants completed the task proposed more successfully with the wide and shallow menu. One can argue that the structures of these menus are too extreme to be completely relevant as far as recommending a menu structure is concerned. Allowing only two items per node on six levels constrains users to make many selections, which is very time consuming. This experiment also revealed that the rate of synthetic speech, the time allowed for user input and the delay before the menus became available had a significant influence on the subjects' performance.

The question of depth versus breadth of menu has been largely addressed, mostly in the case of full-screen visual menus. When depth is emphasised the number of options available in each menu is minimised, whereas the broad menus minimise the number of menus traversed while increasing the number of options available in each menu. In general, results indicate that broad and shallow menus are superior to deep and narrow menus, both in terms of accuracy and execution time [Kig84, PHSV88]. In contrast to these results, designers of strictly auditory menus suggest that auditory menus should never have more than three options at any level because of working memory limitations, thus forcing a deep and narrow structure [ER89, Pel89, GB84]. However the top-level menu option labels of a deep and narrow menu structure will tend to be less semantically similar to the structure's terminal options than will the top-level options of a broad and shallow menu structure for the same menu domain. Huguenard *et al.*'s study on working-memory showed that, contrarily to guidelines for the design of TBIs, deep menu hierarchies (no more than 3 options per menu) do not reduce working-memory error rates in TBIs [HLJ$^+$97].

As mentioned earlier, one critical aspect of navigation in an auditory interface is working memory (because information is presented sequentially). Virzi addressed this issue in an experiment: he undertook three different menu designs, which coincide, from the user's view point, with three distinct dialling plans[3]. The dialling plans investigated were [Vir91]:

4-digit plan: Each category begins with a multiple of 1000, and the last 4-digit number represents an item within the category. For instance, 4009 would be the ninth item of the fourth category.

2+2 digit: The user dials a 2 digit number(10,20,30 ...) to access the category. Then the user dials a second 2 digit number to access the item desired.

Mnemonic plan: Each item is represented by 4 characters. The first 2 characters are letters that represent categories. The last two characters are numbers that represent the items within the category.

The results of this study showed that mnemonic plans were easier to remember but mostly disliked since letters were hard to find on the keypad. These results show that it is easier to build a mental model of a menu hierarchy by associating it to menu names, rather than as an abstract data structure. Consequently, a possible way to improve navigation in hierarchical

[3]A dialling plan is a sequence of numbers that a user needs to dial on a telephone keypad to complete a task.

menus involves helping users to build a mental representation of the menu as a whole rather that as an abstract hierarchical structure.

The question of hierarchical menu structures has been addressed concerning abstract menu structures in the above review. In practice, the use of a menu is very dependent on the semantic content of the menu *e.g.* the menu item labels[4]. Schwartz and Schwab demonstrated the importance of this aspect of the menu design through a study that investigated naming strategies for menu items. Their study consisted of making users regroup menu items in categories in a sorting task and cluster analysis was used to devise the menu categories [SS93]. The outcome of the study showed that an improvement in the classification of menu items could result in an improvement of performance in a telephone-based task. Schwartz and Hardzinsky also suggest that the frequency of use of menu items may also be used profitably to determine the order of menu items [SH93]. The frequency of use of menu items can be involved in the design in two main ways: the first way consists of determining the frequency figures prior to the menu design on a sample of users, and to order the menu items accordingly; the second approach (which can be used as an alternative or a complement to the first one) consists of adapting the menu display according to the user's usage patterns. The drawback of the second approach is the dynamism of the menu structure, which users may find confusing, whereas the drawback of the first approach is that the frequencies do not correspond to the actual usage patterns of the user.

In a case study, Warren designed a telephone menu for a help desk. Warren highlighted three key-points in the design of easy to use menus, which summarise the recommendations made above [War92]:

- The number of choices on the menu should be minimal while still offering the most requested information. People are irritated by being overwhelmed by too many choices.
- Choices should be offered in the order they are most frequently asked for.
- Organise the information so that information common to several choices is only given once.

[4] To be completely accurate, the label of a menu item is only one aspect of the semantics of the menu item. If the item is associated with a picture, the picture contributes to the semantic content of the item. By the same token, if a sound were associated to a menu item, it would also contribute to the semantics of the item.

This section has proposed an overview of research investigating how to design usable menu structures. These studies provide useful guidelines regarding the design of both content and structure of the menu.

2.2.2 Dialogue Design

The previous section, Menu Design focussed on the best possible layout for menu items in a hierarchy. The menu dialogue involves the transitions between these menu items *e.g.*, the means allocated to the user to navigate in the menu and the feedback they receive from the system. The dialogue is the heart of the interaction between the user and the system and, as Aucella and Ehrlich point out, the dialogue flow is critical to a TBI's success [AE86].

As the dialogue design is a fundamental aspect of menu design, it is worth providing the reader with some general background information on the issue. Three main approaches to the dialogue design have been identified in the relevant literature. The first approach, by Schwartz and Hardzinski, envisages the design of the dialogue from the input/output point of view [SH93]. Halstead-Nussolch considers the design with the user's role and the system's role in the dialogue [HN89]. Last, Resnick and Virzi provide a multidimensional dialogue design framework [RV95]. The latter approach, which describes the dialogue design in a more generic way (*e.g.* a way that can be easily extended to all the interaction categories presented in Section 2.1.2), is presented in the remainder of the present section.

Resnick and Virzi envisage TBI dialogue design from the system viewpoint. Subsequently they point out key human factors issues which have to be addressed while carrying out the design of the system. Figure 2.1 shows a standard minimal menu style that does not offer much control to the user. On the other hand, Figure 2.2 shows a style of menu that allows users to skip through the menu items. Navigation is operated by the means of two buttons. This means that menu items are only available one at the time. Note that these examples apply directly to the first interaction category, but the same dialogue description framework could be used to represent dialogues for other interaction categories.

The examples proposed on Figure 2.1 and Figure 2.2 suggest that the design of menu dialogues can be controlled by several parameters. Two of these parameters form a bi-dimensional design sub-space. Such sub-spaces are useful in considering various possible designs. Figure 2.3 shows such a sub-space for the two following dimensions: how to select a menu, and how to advance

Welcome to the ABC Bank's bank-by-phone.

For account balances, press 1.

To transfer money between accounts, press 2.

For mortgage rates, press 3.

To open a new account, press 4. ^

This menu will now repeat. Make your selection at any time.

[presses 4]

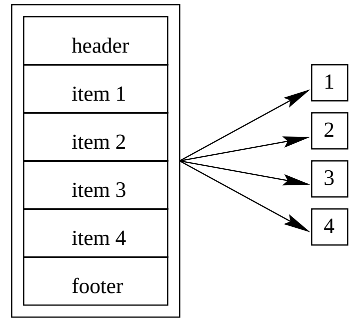


Figure 2.1: Standard menu style. System prompts appear in plain text, while user actions are bracketed. The caret symbol ˆ , inserted in the prompt, indicates timing of the user's key press. That is, the caller presses 4 before hearing that the menu will repeat. On the figure, the enclosing box indicates that the four menu choices are available at any time.

to the next menu. Resnick and Virzi also propose a higher-level framework for the design of TBIs menus. Any design decision is a point in the multidimensional space whose axes are : movement actions, state change actions, action combinations, and user inaction effect (see Table 2.1). The latter representation provides a good reference to dialogue design, which will be needed in forthcoming chapters.

As mentioned earlier, the latter approach by Resnick and Virzi involves TBI menu design from the system point of view. Human factors principles, however, need to be taken into account so that designers make the right choices among the available design possibilities. Resnick and Virzi suggest taking these design decisions with reference to the characteristics of users. Users can be represented on two axes, one representing their familiarity with the content of the system and the other representing their familiarity with the dialogue mechanism (see Figure 2.4). Resnick and Virzi also propose eight design criteria to help designers meet the user's need [RV95]:

DC1: Separate → Flexible A composite action is simpler for users, while separate actions give them more flexibility.

DC2: Positional → Learnable Positional actions are easier to learn, because they are independent of context.

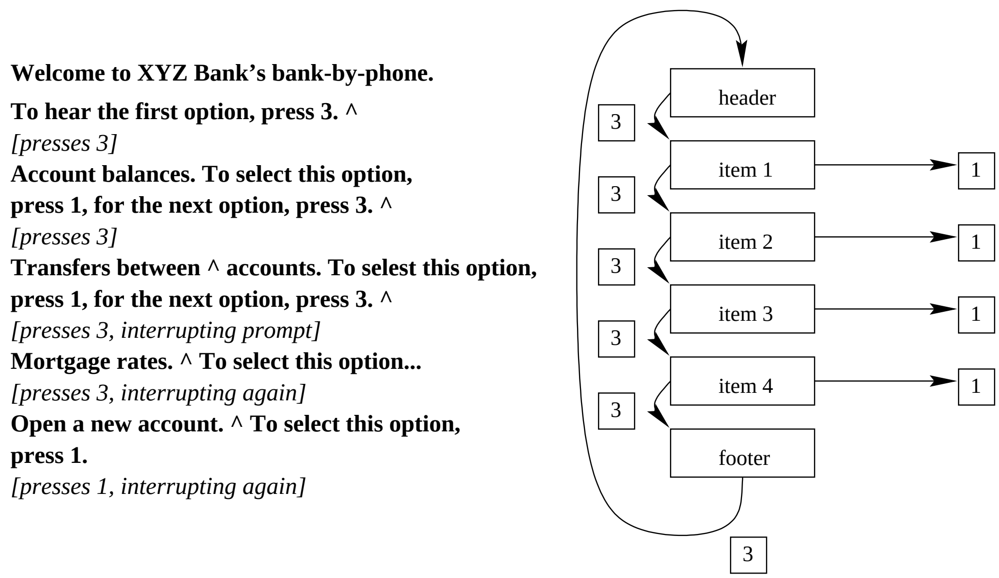


Figure 2.2: A sample two-button menu dialogue. Like in Figure 2.1, the caret symbol ˆ , inserted in the prompt, indicates timing of the user's key press. As the non-adjacent boxes on the figure show, the menu does not go through the successive items automatically (unlike in Figure 2.1). The current item will keep playing until the user presses a button.

DC3: Absolute → Executive Absolute actions are easier to execute than positional actions, once learned.

DC4: Time-out movement a crutch Automatic transitions help mechanism novices, but delay their acquisition of mechanism expertise.

DC5: Moving targets Automatic transitions, together with any position-sensitive actions, create a *moving target* problem for all users.

DC6: Time penalties When information relevant to a user is preceded in a recording by irrelevant information, that user will pay a time penalty.

DC7: Extra prompts → Learnable Describing optional mechanisms degrades usability for mechanism novices, but encourages them to become experts.

DC8: Statements → Passivity Questions, commands, and pauses encourage users to take action right away. Users are more likely to wait for additional instructions when they hear a statement.

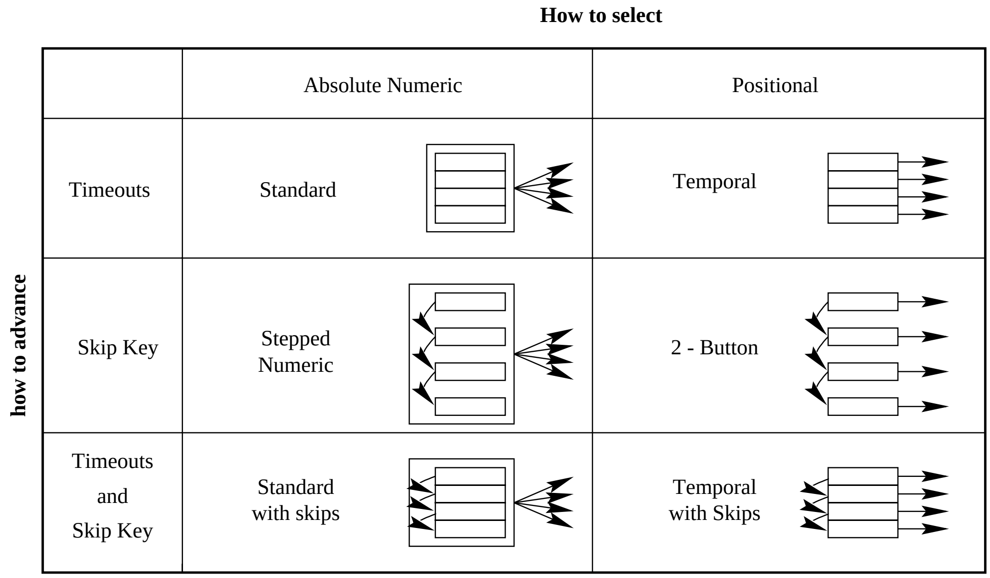


Figure 2.3: Menu dialogue design sub-space defined by two dimensions: how users select items and how they advance from one item to the next.

2.3 Command-Based Interfaces

In a user interface, commands are generally useful for a specific category of users to accomplish specific tasks, but they do not eliminate the need to present lists and menus on occasion. However, as seen in Section 2.1.3, the fact that auditory interaction is sequential makes menus slower to use than in a visual context. Consequently, commands offer an excellent alternative to menu navigation as far as usage time is concerned.

Command-based interfaces typically break down the constraining temporal serialism of menu-based interfaces and substitute it for a human-machine conversation. In addition to being a quicker means to interact in non-visual conditions, it is also possibly more natural [Yan94, YLM95, Yan97]. A lot of research is now dedicated to the development of this dialogue paradigm, for all sorts of applications [MS96b, MS96a, SASH93, GWM98, WC98a]. This corpus of research suggests that command-based interfaces are the interfaces of the future for handheld devices with limited visual feedback.

However, the "naturalness" of dialogue-based interaction should not be taken for granted. What is a natural interface anyway? The main aim of designers must be that interacting with an

Design Dimensions	**Possible Values**
Movement actions	next, previous, repeat, go to header, forward, back, goto.
State change actions	select (absolute), select (current), deselect (absolute), deselect (current), deselect (all), terminate, select (current) + terminate, select (absolute) + terminate, deselect (all) + terminate, delete (current), insert (current location or at the end), assign label to current item.
Action combinations	select/deselect + next.
User inaction effect	Any movement, state change or combination action

Table 2.1: Dimensions of the design of a TBI menu (from [RV95]).

interface seems flawless so that users *feel* that interaction is natural or intuitive. The following question arises: is it reasonable to aim at making interaction between a person and a machine as similar as possible as communication between two individuals? Chin addressed the issue in a user study, he drew up a list of traits that users would like to recognise in interactive voice responses. Respondents selected 45 different traits as desirable for voice mail systems. The top five traits are: practical, intelligent, courteous, efficient and straight-forward [Chi96]. These results clearly make sense within a completely anthropomorphic design framework. However there is no evidence of the fact that human-machine communication can become as natural as human-human communication, even if the machine properties come closer to human properties. Schumacher and colleagues consider that, on the contrary, there is no need for the interaction between a TBI and a person to conform to the rules of conversation between two people. In fact, "creating a telephone-based interface dialogue that rests on human conversation leads to inefficiency and confusion" [SHS95]. In the latter study, by Schumacher and colleagues, command-based dialogues are recommended to be used on three particular occasions:

- The interactive voice response system is used frequently.

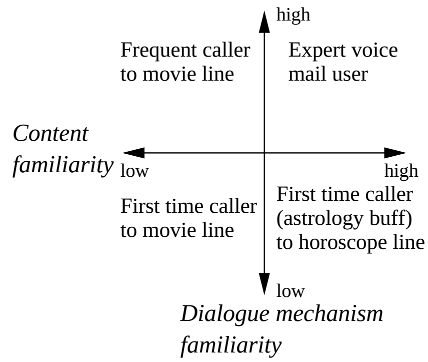


Figure 2.4: Classification of the users in four categories.

- The users are expected to be experts.
- The users need to change modes frequently.

In addition to these interaction issues, one should not neglect the social aspects involved in using conversation-based systems. As Gaver points out, there are several concerns to the use of this technology [Gav97]: The main problem is that audio is inherently a public medium. It is possible to keep audio output private using headphones, but it is more difficult to keep the audio input private too.

Assuming that the aim of dialogue-based interface designers is to reproduce the human to human communication situation; the degree of resemblance achieved is critical to the success of the interface. Current technology can cause the dialogue flow to be easily broken. For instance, silences of the system can be difficult to interpret and can easily become misleading [WC98b]. A brief silence on the system's side can equally mean that it is listening, that it is processing or even that it has crashed. One might argue that in human communication it can also be difficult to tell whether one's interlocutor is "listening" or "processing", but in fact body language, and a lot of other features of human communication generally make speaking with a human a much more enjoyable experience than speaking to a machine.

The issues involved in the development of effective conversational systems go far beyond the development of effective speech-recognition engines. These issues involve making machines intelligent enough to speak. Until then (and even beyond), there is a strong argument as to why menu-based interface will remain a very popular and effective way to interact with a machine, even in a non-visual context.

2.4 Navigation

The review of interaction problems with TBIs has demonstrated that navigation in hierarchical menus is an issue, especially for interfaces with limited visual feedback. The previous section has shown that conversational systems might not be the ideal solutions for this problem. The objective of the present section is to shed light on what navigation really involves and to use this knowledge as a basis to support navigation in forthcoming chapters of this thesis.

The problem with the term navigation is that it involves a whole series of actions. Browsing, scanning, searching and so on, are all related to navigation, even though these are distinct concepts. Therefore, it is necessary to clarify what the term navigation will refer to in the rest of this thesis.

The present introduction to the concept of navigation is followed by a review of the principal techniques developed for navigation in various kinds of digital spaces. This section will then conclude with an examination of the main subtasks and strategies underlying navigation in information structures such as hierarchical menus. The latter examination will set the foundations for the introduction of non-speech sound at the appropriate stages of the navigation process later in this thesis.

2.4.1 Navigation Techniques

The most significant progress regarding navigation in the digital world has come from direct manipulation, on which the desktop navigational framework is founded. As for the manipulation of large sets of data, the direct-manipulation paradigm has been investigated further, with the development of a number of techniques. Beard and Walker have implemented such techniques, like zooming and using a wire-frame box to aid users remember their location in a 2D-space [BW90]. Their study demonstrates that the latter techniques allows users to navigate quicker through the information space than they would with a scrollbar. The study also proved the zooming and framing techniques less efficient than the use of a map window in which the whole structure of the space is available. This point indicates that no matter how good a novel navigational technique is, it cannot match the presentation of the whole information space.

The idea that a persistent representation of the whole space is the ultimate navigational aid should however be moderated for two main reasons: first of all, an information map presents, by nature, a global representation of an information space. Such a representation might not necessarily provide the user with the desired information, which is generally: how to make a transition between one point of the space to another. Geographical maps provide a real-world example of this point; a geographical map is a global representation of an actual geographical area. It is arguable that maps of this type are not necessarily the optimal means to find one's way. Often, what is needed to go from one point to another is a sequential list of moves to make, from one's initial location to the destination. Secondly, information maps are *representations* of the information space: users are not provided with information about the space as they perceive it, but as it can be represented. Again, road maps are a perfect illustration of this: the fact that these maps are only paper representations of the physical world can make it difficult to identify a road.

Even if information maps are not the optimal solution to navigation problems in information spaces, it is arguable that they are a good one. Unfortunately they take a lot of space. As a result, information maps can hardly be implemented in devices with limited graphical displays, such as TBIs or hand-held devices.

The development of navigation techniques consists in finding the right balance between providing information about the global structure of the information space, and providing information about the user's focus point in the information space. One can argue that this task involves the three following points:

- Optimal viewpoint over the relevant points of focus of the information space.
- Optimal means to switch between the relevant points of focus.
- Optimal representation medium of the information.

The *fish-eye view* illustrates this quest for the right balance between global information and focused information (see Figure 2.5): In a fish-eye view, the whole information space is visible at once and it is easy to set the focus on any point of the space (see article by Furnas for general information about fish-eye views [Fur86] [5]).

[5]The interested reader should also consult a comparative study with other interaction techniques, by Kaptelinin [Kap95], including scroll bars and map windows [BW90, ES90]

Mukherjea and Foley's work illustrates the search for effective interaction techniques: they claim that, in that area, navigational techniques are really effective if they show, not only the link and node topology, but more information about the underlying information space [MF94].

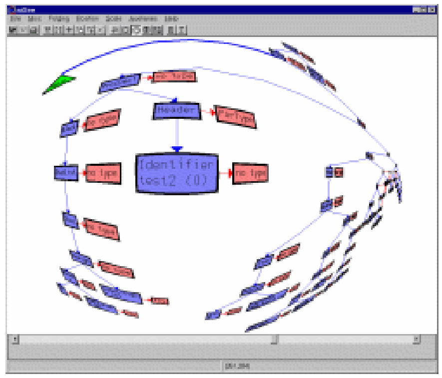


Figure 2.5: Representation of a graph in fish-eye view using aiSee (http://www.absint.de/aisee).

The interesting aspect of the presentation of various navigation techniques described in this section is the goal each of these techniques tries to achieve. Indeed, whether it is allowing users to navigate more rapidly in an information space, or providing them with a better overview of the space, these goals provide insightful information regarding various aspects of navigation.

2.4.2 Navigation in the Information Structure

In the introduction of this section, it has been shown that one of the problems related to dealing with navigation is the generality of the term. Canter's classification is helpful to understand what navigation really involves. Canter proposes a categorisation in terms of the *styles* and *strategies* used by users to navigate in a menu structure [CRS85]. Canter introduces six indices to categorise the movement of users in a hierarchical structure:

- *Pathiness*

  A path is any route through the data that does not visit any node twice. It starts at one point and terminates at another. Menu traversal may be characterised by many short paths (high pathiness) or few long paths (low pathiness).

- *Ringiness*

  A ring is a route through the data that returns to the node from which it started. Since a ring has a home base, it may be thought of as an "outing". Such a ring may include subordinate rings. Menu traversal may be characterised by many rings returning to home base (high ringiness) or a few rings (low ringiness).

- *Loopiness*

  A loop is a ring which contains no other rings. A loop is a simple ring and is distinguished by the fact that no node is visited twice except the home base.

- *Spikiness*

  A spike is a route through the data which goes out to a node and returns exactly the way it came. Hierarchical databases are likely to result in high spikiness since one could traverse the hierarchy down and retrace the path back out.

- *NV/NT*

  The ratio of the number of nodes visited (NV) to the total number of nodes available in the system (NT) give the proportion of available nodes utilised by the user. A high NV/NT ratio indicates a more comprehensive coverage of the menu.

- *NV/NS*

  The ratio of the number of different nodes visited (NV) to the total number of visits to nodes (NS) gives the proportion of first time visits. A low NV/NS ratio indicates a high degree of repetitive visits to nodes.

The indices that best suit the trajectories in a hierarchical structure depend on the user's task and strategy towards accomplishing that task. Canter proposes five very useful indices to characterise search strategies [CRS85]:

- *Scanning*

  When users are scanning, they tend to cover a large area of the menu system, but without going into great depth. Scanning will result in long spikes and short loops which traverse through the database, but do not extend very far into it. It is characterised by a high proportion of nodes visited relative to the total number of nodes available.

- *Browsing*

  Users may be happy to go wherever the data leads them. Users will pursue a path as long as it sustains their interest. Browsing behaviour can be characterised by many long loops and a few large rings.

- *Searching*

  When users are searching for a particular target, the pattern may include ever-increasing spikes with a few loops. It is also characterised by a high redundancy of nodes revisited relative to the total number of different nodes visited.

- *Exploring*

  Many different paths of medium or short length suggest that the users are trying to grasp the extent and nature of the database. They may be attempting to gain a global map of the menu system.

- *Wandering*

  Users may wander more or less randomly through the database. The unstructured journey will lead to many medium-sized rings.

In addition to the search styles and strategies, one needs to understand how the information space is processed mentally by users and on what bases users make their decisions. Norman's models account for the mental process involved in making navigational decisions in two fundamental cases [Nor91] where the user knows explicitly where the item he/she is looking for is located and where the user does not explicitly know where the item he/she are looking for is located. The mental processes involved in these two cases are described respectively in Figures 2.6 and 2.7.

Figures 2.6 and 2.7 illustrate how menu navigation is highly semantically driven, that is, it relies on the menu items labels. This fundamental feature of navigation will be utilised in Section 5.2 as one of the pillars of the framework proposed in the present thesis. However, navigation is not solely semantically driven; people also use *contextual* information as navigation cues. Indeed in a study focusing on navigation in hierarchical structures, Howes points out three features of the human behaviour involved in menu acquisition: namely learning devices by exploration; improving with practice and acquiring display-based knowledge [How94]. Howes' second point illustrates how navigation is more than semantically driven. This suggests that, in a desktop menu, for example, people use elements of the graphical display to learn the structure of the menu. It has been shown that when in front of a device, users could effectively perform a succession of menu selections, but away from the device they could not report the names or order of the constituent commands used. In other words, users do not remember complete streams of actions, but rather build a mental model of the menu structure in which, at each

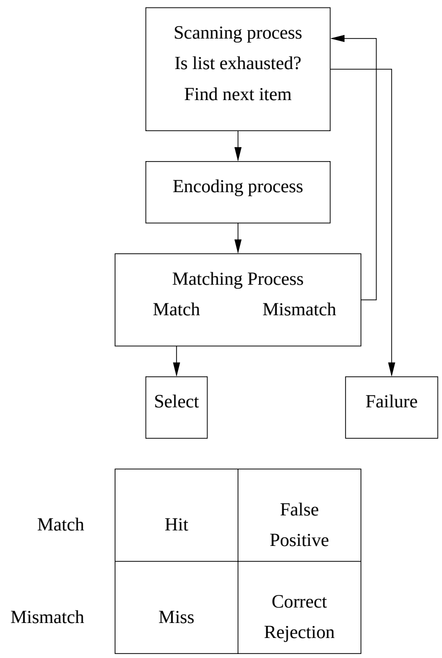


Figure 2.6: Mental process involved in a search task when the target is explicitly known by the user (from [Nor91]).

point of the stream, they know what action to take next to progress towards their goal. In this process, the context in which each action is taken helps recollection of what the next action should be. The visual context in which navigation takes place plays a significant role in this process. As a consequence, it is important to look for alternative methods of providing navigational cues in devices with restricted graphical display.

2.5 Conclusion

This chapter has presented the main issues related to interaction with restricted display devices, with a focus on TBIs. This review has showed that some research had lead to the creation of guidelines for the design of TBIs. Schwartz and Hardzinski have compiled guidelines that exhibit a noticeable effort of communication and standardisation [SH93]. However, these guidelines as well as others [HN89, SHS95] leave a number of unresolved issues.

In particular, the issue of navigation in large hierarchical TBIs' menus remains. If conversation-based systems seem like an interesting alternative to menu navigation, there are still instances

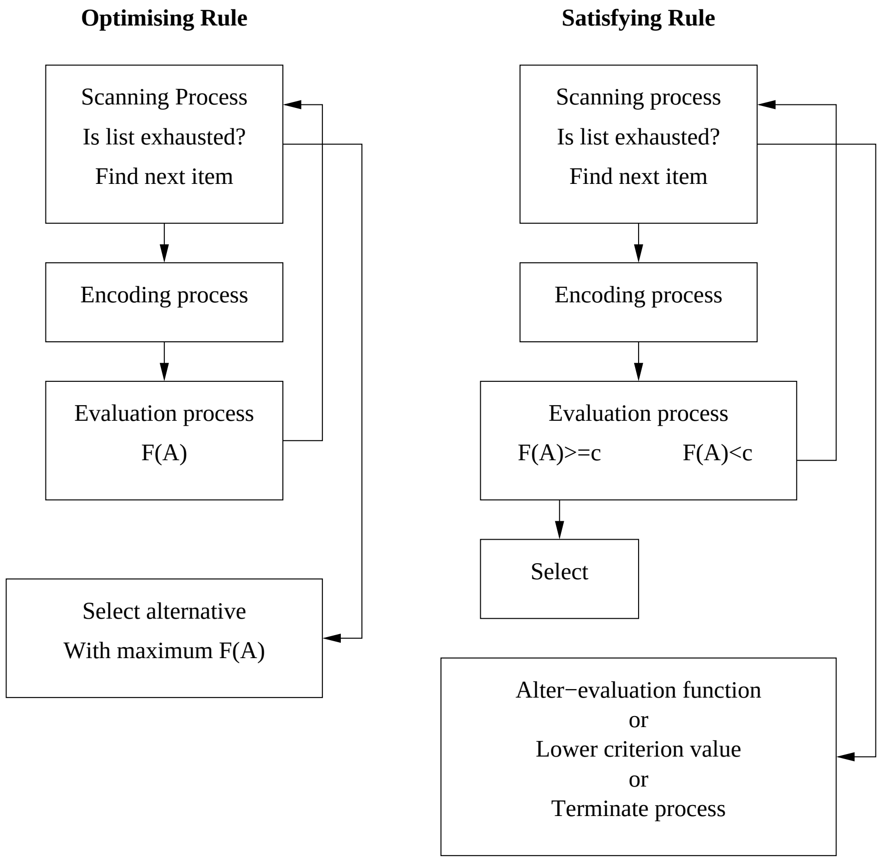


Figure 2.7: Mental process involved in a search task when the target is not explicitly known by the user (from [Nor91]).

in which menus are preferable to commands.

There are two chief reasons why the auditory channel should be considered as a means to improve navigation in limited-display devices. Firstly the visual channel, limited by the size of the devices' screens, can hardly convey more information. Secondly, non-speech sounds have shown a potential for supporting navigation in hierarchical menus (see next chapter for more details and references). The next chapter will introduce the issues involved in designing sounds in a meaningful fashion in order to make progress in answering the following questions:

- Is it possible to design sounds for a large and complex hierarchical menu and how?
- How should sounds be designed to support navigation tasks?

Chapter 3

Perception of Sound and the Use of Sound at the Human-Computer Interface

3.1 Introduction

Cognitive science has long provided grounding for the design of human-machine interfaces. The specific area of auditory interface design is no exception. Psychoacoustics[1], has to be used by researchers in auditory display[2] in order to make use of the constitutive parameters of sound in a perceptually meaningful fashion. The auditory world is large and its understanding pertains to different areas of research. As a significant part of this thesis involves providing sound design principles, it is necessary to understand the mechanisms involved in the perception of sounds.

The first step in understanding the perception of sounds involves understanding sound itself. This chapter will give a brief overview of the physics of sound prior to discussing the most important aspects of the perception of sound. Understanding the perception of sound is related to understanding the nature of sound. What is the nature of the sounds involved in the present thesis? Are they musical sounds, structured non-speech sounds, or simply non-speech sounds? And do sounds have a “nature” anyway? Consider the example of the sound of a church bell.

[1]Psychoacoustics is the scientific field that studies how we perceive sounds.

[2]Auditory display is the field that investigates the the use of sound to represent information.

Such a sound used in a musical piece would be qualified as being a musical sound, for instance, because of the interaction between its pitch and the harmonics of its spectrum, with the other elements of the piece. When one counts the number of times the bell tolls to find out the time, this same sound would not be described as musical. This example illustrates how the perception and categorisation of a sound depends on the context in which the sound is heard or listened to.

Taking a sound from our environment and integrating it into a piece of music has been common practice in the musical repertoire of the past decade. This thesis does the contrary, it uses so-called musical sounds and integrates them in an auditory environment; or more specifically, into the auditory environment of someone who uses a TBI. This process requires a thorough understanding of the perception of "isolated" sounds, as well as the perception of organised sounds (sounds organised in a musical fashion for example). Ideally, some knowledge of the perception of sounds depending on the listener's task would be useful but, unfortunately, such results are virtually non-existent.

There are two main ways to envisage the perception of sounds in a user interface: bottom-up, or top-down. The former approach involves the mainstream cognitive school whose results will be used throughout this thesis. The latter is related to Gibson's ecological approach to perception [Gib79]. Gibson's approach has been applied to the perception of auditory events and provides a lot of information regarding the perception of sounds in our environment, as opposed to, in a laboratory [Gav93b]. One of the main aspects of Gibsonain theory is that animals perceive affordances *i.e.*, a relationship between them and the environment. For instance, when we perceive a chair, we do not perceive its height, but what the chair affords us *e.g.*, whether we can sit on it or not, whether it looks comfortable or not. This approach is therefore extremely useful in auditory design to convey affordances of a system to the user. However, Section 3.6.3 will demonstrate that this approach can not be used to tackle the problem investigated in this thesis. Consequently, the psychological review proposed in the present chapter will focus on the approach which will retrospectively prove relevant later in this thesis *i.e.*, psychoacoustics and music cognition.

On the one hand, psychoacoustics provides insight concerning the parameters involved in the perception of sounds, regardless of their aesthetic qualities. On the other hand, music cognition is concerned with the perception of the musical attributes of sound. The distinction between these two aspects of sound is not always obvious. Section 3.3 will provide helpful information concerning issues pertaining to both these areas. This introduction to the psychology of sound, and musical sounds in particular, will help assess to what extent research carried out in these areas can be exploited in the present thesis. The following sections will then be dedicated to the presentation of relevant studies in the field of sound perception which will serve as a reference for the sound design principles introduced in Chapter 5.

3.2 The Physics of Sound

An *audio source* is a vibrating object whose vibration creates a perturbation that propagates in an elastic medium and reaches our ears to create a sensation referred to as a sound. As a consequence a sound can be represented by a number of sine waves of various frequency and amplitude that vary over time. Each of these sine waves correspond to a *partial* of the sound. The group of partials of a sound define the *spectrum* of the sound. The spectrum is the main component of the attribute of sound called *timbre*, but timbre involves more than just spectrum[3]. When the partials of a sound are all multiples of a common frequency, the sound is called *harmonic*, and the greatest common denominator of these partials is referred to as the *pitch* of that sound. Most of the sounds we hear are rarely purely harmonic sounds. A sound that only contains one partial is called a *pure tone*. These sounds are very uncommon in nature but are widely used in experimental psychology because they are easy to generate and to control. The intensity of a sound is commonly measured in *decibels* (dB) and the perceptive sensation of intensity is called *loudness*. All the components of a sound vary over the duration of the sound. This variation can be represented by temporal envelopes. A global envelope is commonly used to represent the average intensity envelope of a sound. As Figure 3.1 shows, this envelope typically involves four stages: *attack*, *decay*, *sustain* and *release*. It is not uncommon to approximate this envelope to three stages: attack, sustain and release. In this case the last stage is sometimes called decay.

3.3 The Realm of Music Psychology

As music psychology has an important application in the present study, this section is dedicated to surveying the historical development of this discipline. The aim of this brief survey

[3] *cf.* Section 3.4.4

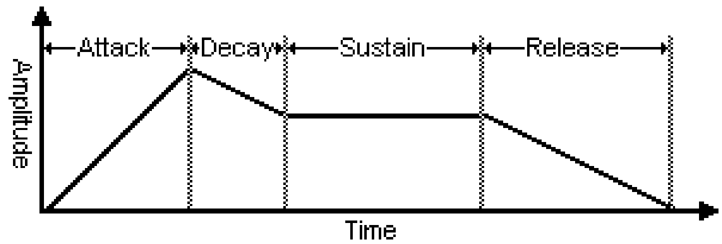


Figure 3.1: Intensity envelope of a sound.

is to shed light on today's conception of sound and the mechanisms involved in its perception. Music perception involves many and varied fields of research as has been identified by McAdams [McA87]:

- Perception of musical qualities of sound.
- Cognitive processes of organisation, representation and storage of music.
- Acquisition and exercise of musical skills in performing and listening.
- Emotional and aesthetic responses to music.
- The nature of inborn and learned musical aptitude.

McAdams also usefully draws a historical portrait of the research trends which have led to today's approaches to the psychology of music [McA87]. For Pythagoras, the perception of sound and music fitted in with the *laws of nature* in that they were related related to the perception of ratios. Accordingly, consonance would result in what he called "pleasure-in-proportion". Pythagoras also assumed that there was a direct correspondence between external or physical dimensions and the resulting internal or perceptual dimensions of sound. This assumption has since been modified by psychophysics, whose aim is to point out the complexity of the link between these specific dimensions.

It is only in the second half of the nineteenth century that the psychology of music departed from purely theoretical considerations and moved into laboratories, becoming experimental psychology. Fechner and Helmholtz can be considered to be the origin of the field of psychoacoustics [Hel68]. The former, in addition to developing a lot of its experimental methods, established the field of experimental aesthetics. In contrast, Helmoltz investigated more general psychoacoustic aspects of musical sounds rather than the musical taste of the individual.

The mentalist school, active between the two World Wars, adhered to the belief that the body and mind were separate, but parallel entities. According to this representation the ear *presents* acoustic information to the great repository of musical talent, the mind (See [Mur37, Sea38, Sch40]). A belief in this clear dichotomy is difficult to maintain in the face of more contemporary evidence that the information received by the sensory systems is known to be progressively integrated at all levels of psychological processing along the sensory pathways. As McAdams emphasises: "The notion that the ear conveys a sensation to the mind which the mind in turn listens to begs the question of how the mind might "listen", and obscures the fact that the coded pattern of brain activity during perception *represents* the musical experience" (in [Spe80]).

For *Gestalt psychologists* there was an isomorphic relationship between the pattern of stimulation and the pattern of nervous activity in the brain [Koh29]. Current knowledge of neurophysiological processes makes this claim untenable, but what does remain from the work of the Gestaltists are two sets of analytical tools which are relevant to the perception of the properties of objects:

- the *figure-ground phenomenon*. This distinguishes some "object" to which we might pay attention (the figure) in spite of the presence of a lot of other information which is relegated by perception to the background (the ground).
- a series of *laws of perceptual organisation*. These attempt to resolve ambiguities about what is figure and what is ground.

The mentalist and Gestalt stances were fiercely attacked by the *behaviourists* who adopted the empiricist principles of the association of ideas and, more specifically, applied them to the association between environmental stimuli and behavioural responses.

A more recent approach to the psychology of music that picks up some of the trends of mentalist and Gestalt theory, is what might be called the *cognitive* approach. Its most important goals include the attempts to understand the nature of mental representation and memory, the processes of organisation of perception and thought, and the ability to reason and solve problems. Sloboda remarks that it is only with the work of Lerdahl and Jackendoff [LJ83] that the psychology of music comes of age [Slo86].

This brief historical introduction shows the evolution of the psychology of sound and music and should provide a context for the results reviewed in the following sections.

3.4 Psychoacoustics

Human perception of sound is an extremely complex matter; pitch does not vary directly according to frequency, certain sounds have the effect of diminishing or obscuring one another to the human listener. Psychoacoustic studies that identify and explain these phenomena are essential to any sound design project. Therefore in this section, physical qualities of sound, and how variations in these qualities are perceived by humans, will be explored in turn, in order to qualify later choices of sound design.

3.4.1 The Audible Field

The human ear can perceive acoustic signals with frequencies ranging from approximately 20Hz to 20kHz. Figure 3.2 (from [ZF90]) shows different thresholds that characterise audition. The bottom line corresponds to the absolute audition threshold below which no sound can be heard. When a subject is exposed to high or continuous acoustic levels, this curve tends to move up. Amplified music concerts, work noises and personal stereos are typically responsible for irreversible *acoustic traumas*. The dotted line on Figure 3.2 (in the bottom right corner) represents the absolute threshold of audition for an individual who has frequently listened to music at high volume during long periods of time. The position and amplitude of these peaks depend on the nature of the acoustic exposure. When this curve reaches the frequency level of speech, this might result in a loss of comprehension in day-to-day acoustic/auditory interaction.

Figure 3.2 clearly shows that at equal acoustic pressure level, the human ear perceives medium and high frequency sounds better than low frequency sounds. Accordingly, low pitch sounds should be used carefully by the designer, as they might not be heard properly. Consequently, from a practical point of view, high pitch sounds are appropriate to applications such as alarms, while lower-pitch sounds may be more suitable to provide continuous feedback in a less obtrusive fashion. This inevitably conditions the use of sounds in the present study.

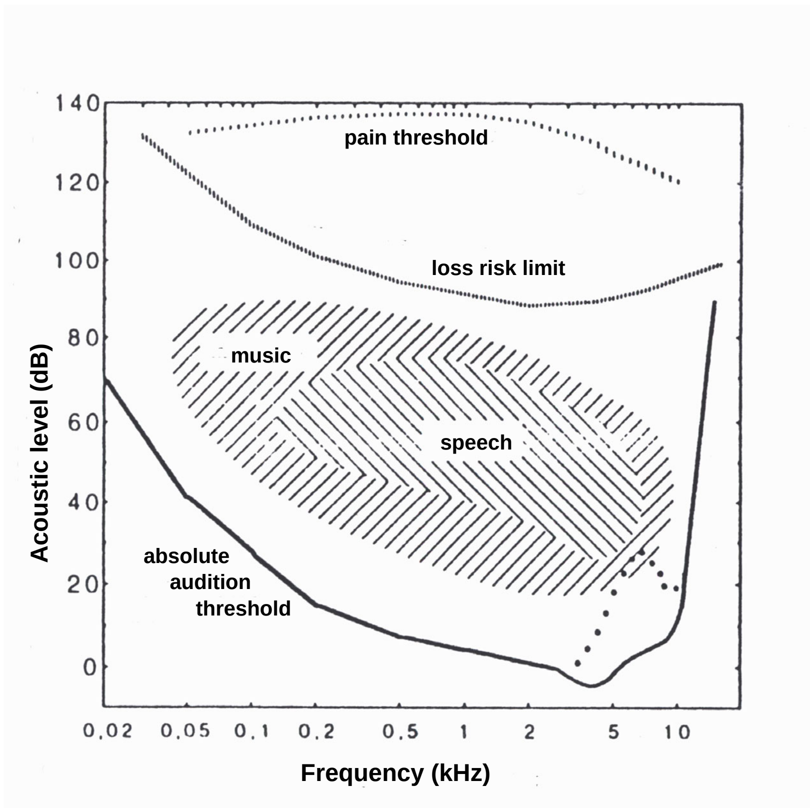


Figure 3.2: Audible field for a young adult with a good audition (From Zwicker and Fash [ZF90]).

3.4.2 The Ear as a Spectrum Analyser

The human ear is capable of analysing complex auditory scenes[4] where sounds come from different sources and allowing the listener to focus on one of these sources. This faculty is limited by two characteristics of our auditory system. The ear cannot segregate sounds whose difference of frequency is below a certain threshold and is also constrained by a phenomenon of masking due to the difference of intensity between sounds. Indeed our audition is subject to a maximum differential in both the frequency and the intensity axes.

Frequency Resolution

A study of the interactions between sounds with close frequencies show the limits of the frequency resolution of our hearing system. For instance, the audition of two pure tones[5], whose frequencies are very close, results in the perception of a single tone with an amplitude that varies. The frequency of these intensity variations depends on the distance between the tone frequencies. Within a range of frequency called the *critical band*, this phenomenon occurs, whereas tones outside the critical band are perceived as separate entities [Sch61, Sch70]. See work by Plomp concerning the interactions occurring within these critical bands of frequencies [Plo76])

In practice, this phenomenon is easily identifiable because it relates to roughness: below a certain frequency interval within a critical band, we perceive an intensity *beat*. Just outside this interval, still within the critical band, we experience a sensation of *roughness*. Roederer has shown that this sensation of roughness can still be experienced to a certain extent outside the critical band [Roe75].

Masking

When the auditory system fails to segregate signals played simultaneously, it is due to the phenomenon of *masking* [Fle40]. This effect can be said to be *total* when a sound of a given loudness cannot be detected in the presence of another sound. The term *partial masking* marks the reduction of the subjective intensity of a sound at a given level while another sound is played simultaneously. In general, the masking effect is more efficient when the masking sound has a lower frequency than the masked sound. This characteristic is related to the excitation

[4] *cf.* section 3.4.5 on the auditory scene analysis.

[5] A pure tone is a sound whose waveform is made of a single sine wave

pattern on a membrane of the inner ear called the basilar membrane[6]. Figure 3.3 (from [Sch64]) illustrates the masking effect in various situations. This figure shows the masking of a pure tone (S) by a narrow band noise (B) depending on the intensity of the noise and of its pitch. The three bottom pictures on the left of the figure represent the excitation patterns of the basilar membrane for three different levels of the narrow band noise, where the pitch of the noise is greater that the pitch of the pure tone. In these three cases, the masking effect (represented by the hashed area) is minimal. In the right column, the frequency of the noise and of the pure tone have been swapped. In this case, the intense noise masks the pure tone completely.

Patterson suggests that a sound should be played at a level between 10dB and 15dB louder that the level of the background in order to avoid any masking effect, but maintain a reasonable volume [Pat82]. Again, this research is of crucial importance in the field of sound design.

3.4.3 Perception of Pitch

The pitch of a sound is defined as the frequency of a pure tone whose pitch subjectively matches that of the test sound [McA97]. The pitch varies according to frequency for pure tones, and almost always with the fundamental frequency for complex periodical sounds. Non-periodical sounds may also create a sensation of pitch.

As with most parameters of sound, the perception of pitch is affected by many other factors. As Buxton *et al.* point out: "It is important to be aware of the myriad interactions between pitch and other attributes of sound when using pitch" [BGB91]. The human ear can distinguish very small differences of pitch. The smallest interval detectable by the human hear is called the *limen*. This difference varies with parameters such as volume and frequency. The frequency difference limen varies from 1Hz at 200Hz to 68Hz at 8kHz for sounds played at 40 dB. This difference seems large but one should bear in mind that pitch actually has a logarithmic relation to frequency. Pitch perception is also greatly affected by age.

Although the human ear is very good at detecting small differences of pitch, the ability to make absolute pitch judgement is a very rare skill. Moore estimates that only 1% of the population have *perfect pitch* [Moo82]. Conversely, the total inability to distinguish very distinct pitches (*tone deafness*) is also very rare. Auditory interface designers should be aware of these user categories, but they can reasonably regard them as pathological cases.

[6] *cf.* work by Pickles for more information on the physiology of the ear [Pic89]

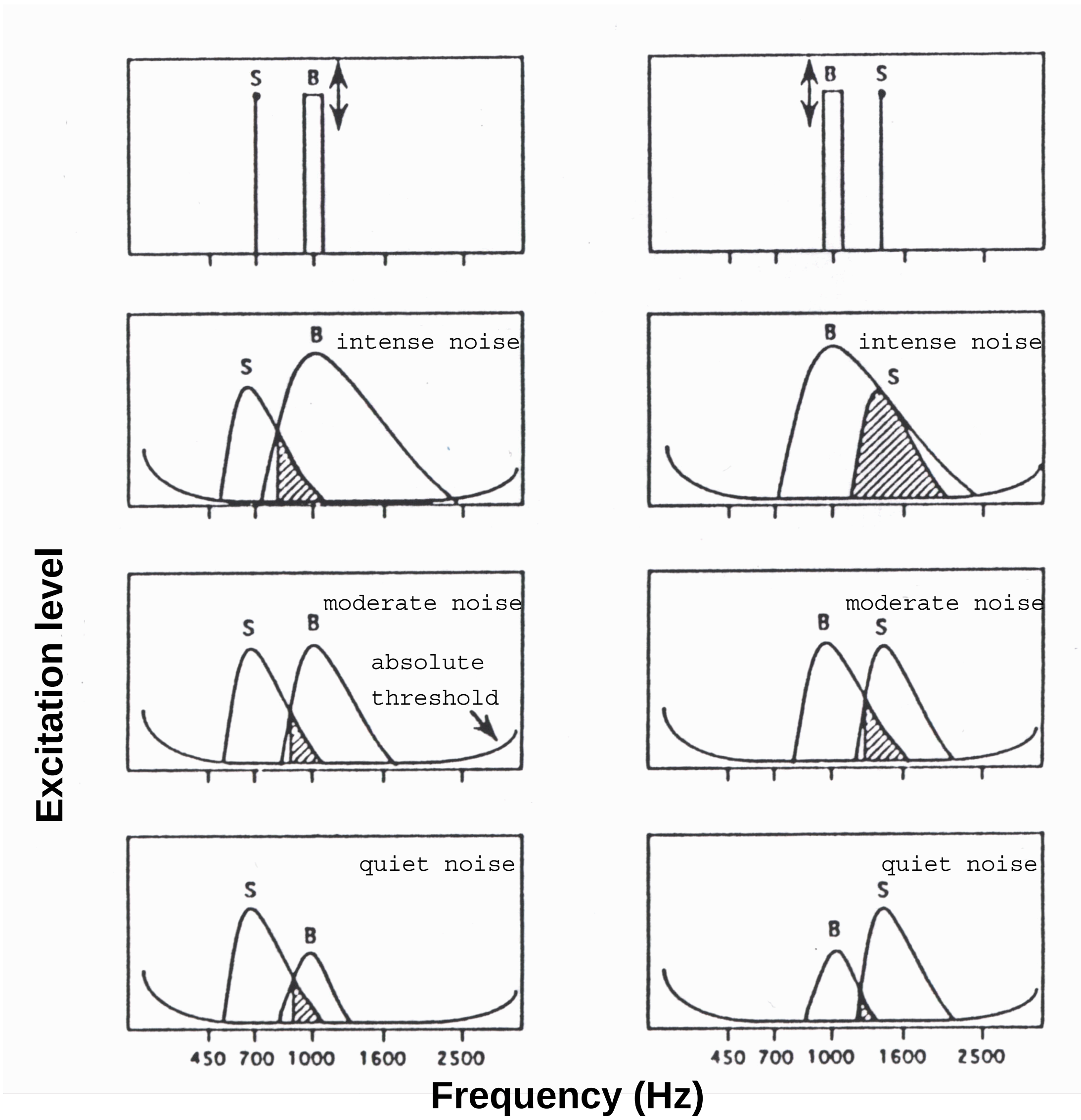


Figure 3.3: Excitation patterns illustrating the masking effect.

3.4.4 Perception of Timbre

Traditionally, timbre is referred to as the *quality* of sound. It is the most interesting and most challenging parameter of sound to study. The National Standards Institute defines it as follows [Ins60]:

> "Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar."

What is timbre?

Timbre as long been believed to be the same as the Fourier spectrum, or even ignored. According to Boring, there are "seven most important" attributes of sound: pitch, loudness, brightness, volume, vocality, tonality and density [Bor42]. As can be seen, timbre is not part of this list. Musical sounds are often considered to be thought of as comprising three sections: attack, steady state, and decay. For Helmoltz, timbre was determined by the spectrum of the steady state. As Risset and Wessel note, this idea creates serious difficulties [RW99]:

> "The saxophone remains a saxophone whether it is heard over a distortion-ridden pocket-sized transistor radio or directly in a concert hall."

In the situations mentioned in this example, the spectra of the sounds are extremely different: Firstly, because they are produced by two different media (acoustically versus through a low bandwidth electronic device) [EE47]. In addition, in a room, the frequency response can fluctuate by up to 20 dB from one corner of the room to another [Wen35]. The lack of identity and liveliness of the early attempts of computer sound synthesis, based on spectra, shows that this factor does not account of the nature of timbre in a convincing way.

This lack of realism has lead researchers to pay more attention to the temporal attributes of sounds. More than its spectrum, it is the temporal evolution of a sound spectrum that matters. Risset and Wessel point out that playing a piano sound backwards gives a non piano-like quality, although the original sound and the reversed one have the same spectra [RW99].

The attack of a sound is a fundamental parameter of sound that affects the perception of sounds and timbre. This is due to the very nature of the auditory system: The auditory system is very good at identifying changes in our environment. The attack of a sound is the primary component of an audio event that allows the auditory system to identify the nature of the event. The problem with attack transients is that they are very brief and complex; they are typically complex non-linear phenomena.

How is timbre perceived?

When designing sounds, the choice of instruments is often a critical one. Therefore it is essential to know whether two different instruments will be distinct enough for users to tell apart. This task is trivial when dealing with instruments that belong to clearly distinct families, such as a clarinet, a violin and a bongo. However, these radical choices of instruments are very prohibiting for a sound designer. In practice the necessity for a sonification[7] to be homogeneous implies that similar instruments have to be used. The use of synthesised sounds which do not emulate any acoustic instrument provides us with a richer variety of timbre, but this also requires more care; indeed such sounds are by essence "new" to the user's ear and it is difficult to anticipate how they will be perceived. This section summarises the factors which influence the perception of timbre.

For ecological reasons, the human ear is very efficient at distinguishing changes of timbres, because it is good at distinguishing changes in the environment in general. However, it is more difficult to assess our ability to categorise and remember timbres.

Timbre is a complex parameter because it is multidimensional and because there are interactions between its dimensions. The multidimensional aspect of timbre began to be systematically studied in the early sixties [Kru64a, Kru64b, She62a, She62b] with the use of multidimensional scaling techniques. Figure 3.4 illustrates how various instruments can be located in a multidimensional space by using multidimensional scaling techniques. This figure represents a spatial model with five latent classes derived from dissimilarity ratings on 18 timbres by 88 subjects. The acoustic correlates of the perceptual dimensions are indicated in parentheses. Shaped lines connect some hybrid timbres (vbn and sbno, respectively an hybrid between a vibraphone and a trombone, and an hybrid between a bowed string instrument and a piano) to their progenitors. For more details on the instruments located in this 3D space, refer to the original article by McAdams and colleagues ([MWD$^+$95]).

[7] See Section 3.6.1 for a definition of sonification

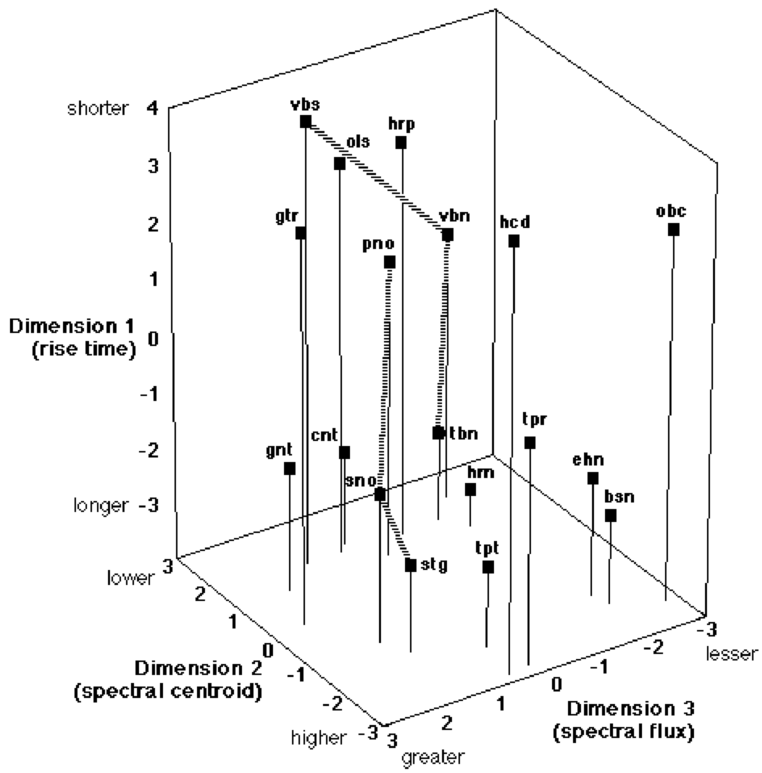


Figure 3.4: Multidimensional representation of a set of timbres. The three representation dimensions of this space are the rise time (Dimension 1, time taken by the instrument to achieve a stationary mode), the spectral centroid (Dimension 2, average frequency value of the spectrum) and, the spectral flux (Dimension 3, temporal evolution of the spectrum), from [MWD$^+$95]

Studies by Wedin and Goude [WG72] have shown that there is a clear relation between the families of musical instruments tones (wind and bowed strings) and the spectral envelope of these tones. Moreover, the results remained unaffected whether the attack part of the tones was presented or not. These results suggest that the spectral characteristics of tones are predominant in human perception of timbres.

Recent research has shown that the temporal and spectrotemporal contributions of sounds is greater than Goude's studies in the perception of timbre suggest. Grey reported that, in addition to the spectral energy distribution, certain additional parameters play an important part in the perception of timbre. These are the presence of synchronicity in the transients of higher harmonics (along with the closely related amount of spectral fluctuation within the tone through time) and the presence of low-amplitude, high-frequency energy in the initial attack segment [Gre76].

Saldanha describes the importance of the attack part of sounds in identifying timbres [SC62]. In an experiment, he presented subjects with 10 instrument sounds in different conditions. The results showed that the part of the sound presented, attack, steady state or decay, had a significant effect on the identification of the instruments. An instrument was best identified when the attack was played. The study also showed that instruments are recognised more easily when played with a vibrato. These results are to be expected according to the very nature of our auditory system which is good at identifying changes in the auditory scene. More recently, Iverson and Krumhansl have confirmed that what matters when a sound is displayed is that the onset of the sound must be clearly heard [IK93]. Wessel also found that similarity judgements were related to the "bite" of the onset of tones [Wes79]. This suggests that the salient attributes of sound have to be carefully considered in the sound design process involved in this thesis. Sounds must often remain brief and therefore there onset is the chief means for these to be identified.

3.4.5 Perception of the Auditory Scene

Now that a better understanding of how the auditory system deals with individual attributes of sound has been dealt with, the perception of more complex *auditory objects* in various contexts can be addressed. This section will explain how a listener perceives auditory information in order to determine the presence, the nature and the position of sound sources of their environment. These issues are obviously of primary importance in the area of sound design, as sound designs are most often made of complex auditory scenes and the designer should know how these scenes are comprehended by listeners. This process is what Bregman calls the *auditory scene analysis* [Bre90]. Figure 3.5 illustrates the latter mental process.

To begin with, the term "auditory object" must be defined as it is fundamental to the comprehension of the perceptive organisation process. Auditory objects are the mental representations of groups of elements that have a coherence between themselves. Such objects are not necessarily categorised by their spatial position, or by the physical properties of the sources that emit them. For instance, a violin sound, a oboe sound and a clarinet sound played simultaneously may not be perceived as three separate objects, but as one single auditory object.

Composers often exploit this phenomenon, creating situations in which auditory sources are combined to form a single musical object, despite being generated by different physical sources (musical instruments). The auditory organisation of a scene at a given time depends on its acoustic content, as well as on the context in which the event occurs. An understanding of the mechanisms involved in the auditory organisation process will help create auditory objects and streams where the user's interpretation is predictable.

There are two major premises regarding auditory organisation [Bre90]:

1. A single acoustic object cannot belong to two or more groups simultaneously.

2. The perceptual qualities of an auditory event are worked out only after the elements of the scene have been grouped into sources *e.g.*, it is the grouping process that determines the perceptual qualities of auditory events.

This is to say that grouping leads to a mental representation in which a source possesses perceptual qualities related to the properties of all the elements of the scene it belongs to. Consequently, qualities such as timbre, pitch, or even loudness depend on the grouping process that happens in the human brain. Ambiguous situations can lead to ambiguous perception of auditory qualities. In the light of recent research results, it seems difficult to give full credit to the first hypothesis of separation. According to Bregman, acoustic objects are not "opaque", as visual objects are, but are "transparent" and it would seem likely that the auditory system allows some sort of multiple categorisation of sources [Bre90].

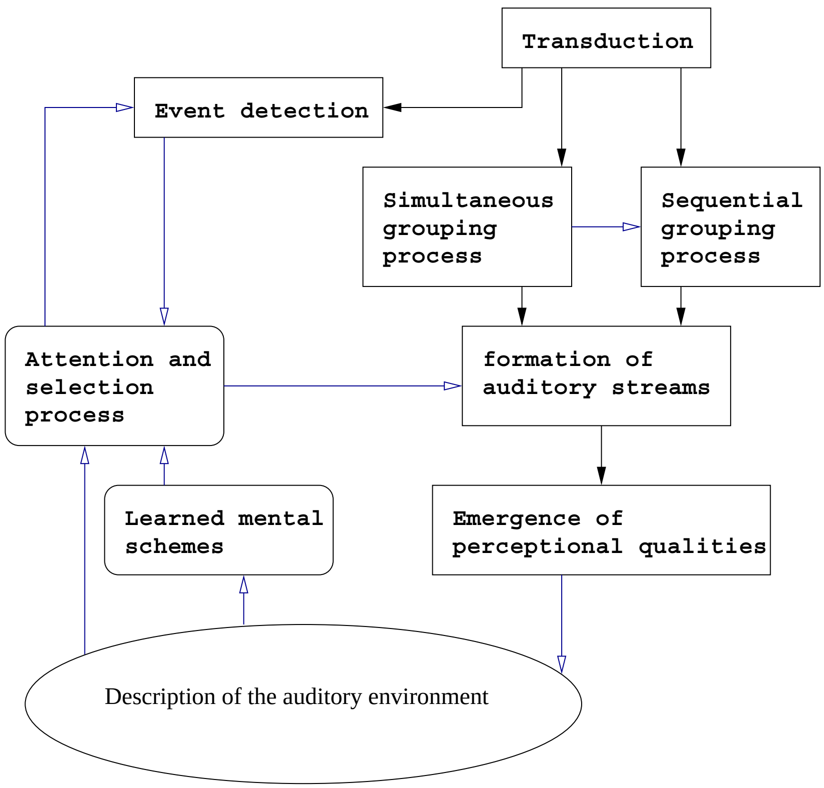


Figure 3.5: Mental process involved in the auditory scene analysis (from [McA97]).

Traditionally, the grouping principles have been devised according to the *Gestalt* theory [Wer38]. The two main processes involved in the mental representation of auditory sources are: a simultaneous grouping process, and a sequential grouping process. These two theories concerning the grouping principles of the auditory system will now be described in turn.

The Simultaneous Grouping Process

This process is used to group simultaneous information analysed by the auditory system from one source, and segregate the information coming from distinct sources. Several cues are used in the process:

- *Synchronicity of attacks and releases*

  The spectral characteristics of a source generally tend to start at the same time [Hel68]. A delay of 30 ms between the attacks of two simultaneous sounds creates a clear decrease of the perceptual fusion of the sounds, and an increase of the categorisation of each sound [BP78, Ras78]. A delay comprised between 40 and 80ms is sufficient to create the impression of a double source, resulting in the perception of two distinct sources [Cut76].

- *Amplitude modulation coherence*

  Some experiments have shown the importance of this factor in extracting a signal immersed in a background noise. The masking threshold of a signal has been measured when the signal is played in a band of noise. The results showed that the masking threshold was lower by 6dB when the amplitude modulations of the noise and of the signal were coherent [HGH88].

- *Harmonicity and separation of the fundamental*

  The harmonicity between spectral relations or a periodicity of the wave forms lead to the perception of a unique pitch in most cases, and subsequently of a unique source. The presence of an in-harmonic series or of several in-harmonic series produces, on the contrary, the impression of different sources [Dem89]. In an auditory scene containing two harmonic series, separation of the fundamental frequencies of at least 3 to 8% is necessary in order to use this information to segregate the sources [Sch83].

- *Spatial location*

  Acoustic objects coming from the same spatial position tend to be grouped together. In order to locate items spatially, the auditory system uses different cues: The temporal difference between the times a signal reaches the ears, the difference of loudness, and the spectral differences between the ears.

The Sequential Grouping Process

The sequential grouping process is used to group successive events which show coherence. This phenomenon has largely been investigated by Bregman [Bre90]. This grouping process operates on the basis of the continuity of the spectral and loudness information of auditory events within a stream.

One of the main properties of sequential organisation is that it is difficult for a listener to make any temporal judgement about separate streams. This notion corresponds to what van Noorden calls *temporal coherence*, which compares against the *fission* of auditory streams [VN75]; by definition a stream must possess a temporal coherence. This phenomenon has been highlighted by experiments such as that illustrated in Figure 3.6. In this situation, depending on the characteristics of the events, the scene can be perceived as two streams or as a single stream. Several parameters are involved in this grouping/segregation process:

- *Distribution of frequency values*

  The main hypothesis concerning this cue is that successive events coming from a same source have coherent frequency distributions. In the case of a simple harmonic sound, this hypothesis means that successive events coming from the same source have coherent pitches *e.g.,* pitches in the same range. Sudden changes in pitch are usually interpreted by the listener as being caused by the presence of other sources. The auditory system establishes perceptual links between events which have spectral similarities. On the contrary, dissimilarities in the frequency distribution of the sounds causes a fission of the sequence in different streams. The degree of separation of these streams depends on the difference between the frequencies of the events, and also on the tempo at which the sequence is played. Higher tempos as well as higher pitch intervals help segregate streams (see Figure 3.6 and the relevant audio samples, Sample 1, 2, 3).

- *Spectral envelope*

  The hypothesis concerning this cue is that sudden changes in the spectral envelope are interpreted as the appearance of a new source. This cue is usually assimilated into the previous one (the distribution of frequency values of the sound), but rigorously, it should be considered separately, the former being related to pitch, whereas the latter is related to timbre. Similar experiments to those described above have shown similar results when using timbres as a grouping/segregation cue. This result shows that differences between instruments, which is a radical difference between timbres, can be used to help listeners segregate streams.

- *Loudness*

  The hypothesis concerning this parameter is that the intensity of a source must vary slowly as compared to the rate at which events occur. Sudden and frequent changes of amplitude could consequently signify the the appearance of a new source. Van Noorden found out that for rates below 2.5 sounds per second, a difference of more than 5 dB between two alternating sounds of identical frequencies is sufficient to produce a segregation [VN75]. As for spectral cues, an increase of tempo and of the difference of loudness helps discriminate the streams.

Sample 1

Sequence of sounds perceived as a single stream.

Sample 2

Sequence of sounds perceived as two distinct streams. The sequence is based on the sequence of Sample 1 in which some of the pitch intervals have been increased.

Sample 3

Sequence of sounds perceived as two distinct streams. This sequence is based on the sequence of Sample 1 in which the notes are played with two distinct instruments.

Interaction between the Simultaneous and Sequential organisation processes

Sequential and simultaneous organisation processes are not independent, they can interact and compete. In general, in real-life situations, there is a balance between the different simultaneous and sequential cues. In laboratories however, artificial conditions in which these cues can compete can be created. What is of importance, as far as designing sounds is concerned, is to have a good understanding of the behaviour of all the individual cues so that conflicting situations are avoided. An important aspect of this chapter is to provide the reader with enough knowledge of the perception of sound in various contexts. Such statements as: "the pitch range between two streams should be maximised to ensure that the streams will be perceived as two distinct streams" are intentionally avoided. The most useful principles are those contained implicitly in the description of the processes involved in the perception of sound. Specific design

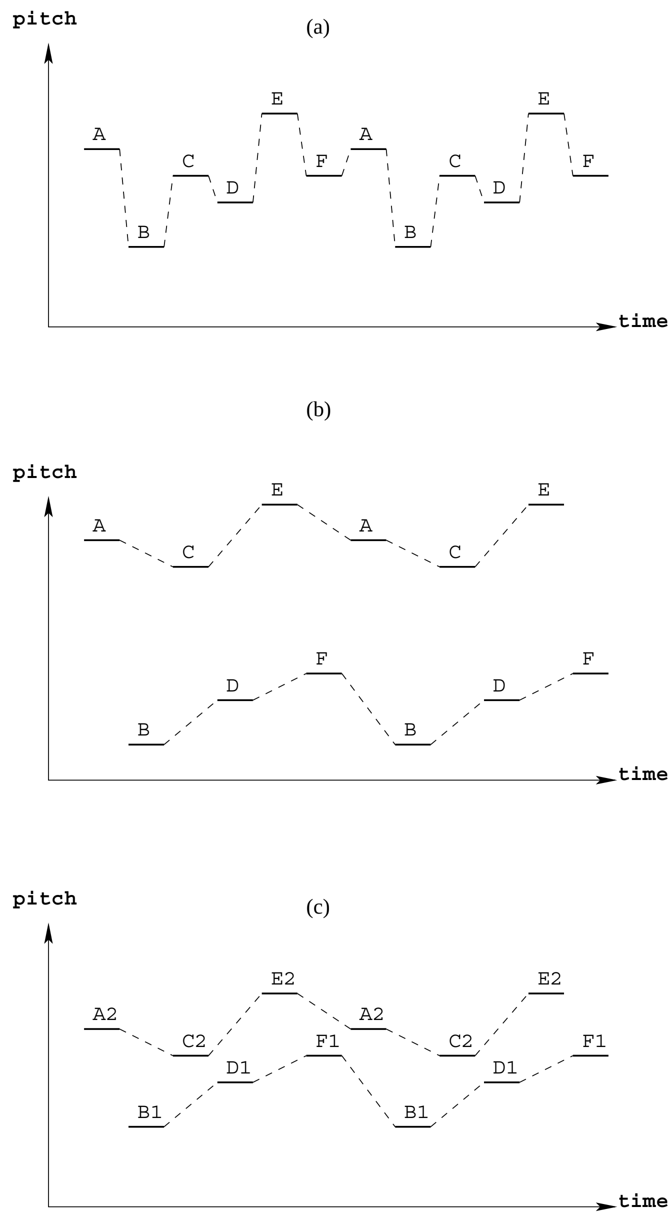


Figure 3.6: Illustration of the grouping/segregation process involved in the perception of an auditory stream. In Figure (a), the listener perceives a unique stream, whereas in Figure (b), two distinct streams are perceived. this effect is achieved by increasing the pitch range between the first stream (Figure a) and the second (figure b). Figure c illustrates how, alternatively, the same effect could be achieved by using two distinct instruments to play sounds A, C, E and, B, D, F, respectively. These three examples correspond to the audio examples: Sample 1, Sample 2 and Sample 3 respectively.

principles will, however, be provided later in this thesis whilst describing the design of sounds to support navigation in hierarchical menus.

3.5 Music Cognition

The current chapter began with a general presentation of music psychology as a science (Section 3.3). After this the main phenomena involved in the perception of auditory events were reviewed (Section 3.4). However, this review was only concerned with auditory events independently from their musical attributes. The aim of the present section is to cover the perception of musical events *e.g.,* auditory events taken in a musical context. The first question that arises is, what makes an auditory event musical? And by the same token, is music organised sound, is organised sound music?

Historically like many forms of art, music used to be created according to strict, well-defined compositional principles. From this point of view, a piece of music is musical by nature. This approach to music has progressively vanished over the twentieth century and nowadays what distinguishes noise from music is simply the listener's judgement. The latter conception regards music as an entity which meets certain aesthetic requirements. In contrast, music is regarded by the former conception as a phenomenon which makes sense universally[8].

The comprehension of music is essentially the result of a learning process and this learning process has two important aspects. On the one hand, there is an *implicit* learning process resulting from the exposure of people to a certain acoustic environment. On the other hand, there is an *explicit* learning process that results from musical education [Fra58]. Consequently, the perception of structured sound is indeed culture-dependent. The point of the present section is to identify the (musical) rules of sound organisation (composition) which do make sense with reference to western musical tradition. These rules are fundamental to the design of abstract sounds.

3.5.1 Pitch Scales

Scales are composed of collections of pitches, these constitute the alphabet of music. In most cultures, scales use from five to seven notes per octave. In western music, the various scales used

[8] Or at least within a given culture.

are subsets of the twelve note *chromatic scale* in which each note is separated from the next by a half-tone. The historic development of scales is a difficult issue as it does not seem to follow the same logic from one culture to another. However Dowling and Harwood have demonstrated the influence of various factors in the development of scales [DH86]. These factors include:

- Humans' ability to discriminate auditory objects (the half-tone is at least six times bigger than the differential frequency threshold). Some cultures however use smaller intervals than the semi-tone.
- The octave equivalence; scales replicate themselves from one interval to the next.
- Limited number of notes per octave.

Scale structures remain relatively stable within cultures and people quickly develop rigid mental auditory patterns *e.g.*, most people are capable of singing the notes of a *major* scale: C, D, E, F, G, A, B, C.

3.5.2 Tonal Hierarchies

In a given key (say C major), the notes of the scale do not have the same perceptional importance *e.g.,* some notes are perceived as being more strongly part of the scales than others. An interesting question is whether the tonal hierarchy taught by music theory is coherent with the hierarchy between the notes of the scale as people perceive it. Krumhansl has shown that both theoretical and cognitive classifications are similar [Kru90]; She measured the concordance of the notes of a given scale with the first degree of that scale[9]. In the case of the major scale, she found out that the best fit was the first degree (tonic), followed by the fifth degree (dominant), followed by the third degree, then the fourth, followed by the sixth. The main hypothesis which explains these results is that the acquisition of a stable tonal representation is related to the statistical properties of the *musical surface* i.e., the occurrences of the various degrees in our musical environment.

Similar patterns between the notes of a scale also exist between the chords of that scale. For example, in C major, the *C-E-G* chord is dominant over the *G-E-D* chord, itself dominant over the *F-A-C* chord. Such hierarchies determine harmonic expectations and reflect the occurrence

[9] The notes of a scale coincide with degrees; in C major, for example, C is the first degree of the scale, D the second, up to B which is the seventh degree. The first degree is also called the *tonic* and the fifth degree the *dominant*.

frequency of various chords within a tonality. Krumhansl highlights that the *intra-tonal* hierarchy can be predicted from the position of the component of the chords in the tonal hierarchy mentioned in the previous paragraph [Kru90].

The first paragraph of the present section has demonstrated that the notes of a scale are ordered hierarchically, and that chords in a given tonality are also perceived in a hierarchical fashion. Tonal music is also made of tonality changes called *modulations*. Again there is an obvious hierarchical structure in the relations between different tonalities. Bigand has shown that a modulation towards a “far” tonality creates a strong rupture in the musical flow and makes the sequence harder to remember [Big91]. This phenomenon is only possible because of an implicit notion of distance between tonalities.

The results of music cognition concerning the perception of tonal relations, reviewed above, are coherent with tonal music theory. The main conclusion to be drawn from this is that the hierarchical relationships of tonal music make sense to musically trained and musically untrained listeners and should be used in sound design. A potential application of tonal hierarchies would be the representation of hierarchical data. For example, in hierarchical data, the most important data could be represented by the first tonal degree, the second most important by the fifth degree of the same scale and the third by the third degree of that scale. Tonal progressions can also be used to represent completion: in the case of a menu sonification, let one assume that a designer wants to allocate a sound to each menu option. He/she may choose a C chord for the first option, a D chord for the second option and a G chord for the third option. The C-D-G-C progression will inform the user that he/she has completed the scan of all the options of the menu.

In musical terms, these results and the latter application example suggest that “cadences” can be used in sound design. Cadences can be defined as an inflection or modulation in tone, or more simply, as any rhythmic flow of sound or the harmonic ending, final trill of a phrase or movement.

3.5.3 Recall of Auditory Motifs

Remembering audio motifs is an important issue for the design of auditory messages that convey information to users. This section will identify those factors that influence the way auditory motifs are recalled.

A lot of research has been devoted to the mental representation of melodies in memory (see [Cro93, DH86] for reviews). A clear principle that comes out of this research is that melodies are memorised in terms of relative pitch intervals rather than in terms of absolute pitch. Using this token, a melody is easily recognised when transposed. The musical context and the intervals involved have an influence on the coding mechanism involved. Concerning the perception of known melodies, Deutsch has shown that minor distortions of a known melody (a few random notes are transposed by one octave) make the perception of the melody very difficult [Deu72]. This finding suggests that, more than the series of notes in the melody, the contour of the melody plays a fundamental role in the recollection of known melodies. Research on recollection of new melodies has identified several parameters:

- The melodic contour.
- The interval patterns and their directions (for example the number of ascending or descending semi-tones between two consecutive notes).
- Rhythmic patterns.
- Absolute pitch values.

To summarise, melodic contour is the main factor in the perception of melodies in the short term and in a non-familiar context (atonal context)[10]. On the other hand in the case of a familiar context (tonal context) and for longer term recollection, pitch interval sizes and the position of notes within a scale are dominant factors.

The main point of this section as far as sound design is concerned is that there is no absolute rule for the design of easy recallable melodies, but relative rules. The recall performance of a melody present in a sonification depends on the other melodies of that sonification. What the above results suggest is that in an atonal sonification, an atonal melody in that sonification will be difficult to recall. However, in a mainly tonal sonification, an atonal melody will be easy to identify although, its contour may be difficult to recall accurately. The latter example may be useful in a sonification of many systems: in a mainly tonal sonification, atonal melodies may be used to represent unusual events such as errors or alarms.

[10] The experiments carried out on the recall of melodic phrases use atonal melodic phrases as non-familiar phrases. This is not to say that a particular atonal phrase can not become a familiar one

In addition, to ensure that the melodies of a sonification are recalled distinctly from one another, the emphasis should be put on designing a large variety of melody types. As pointed out above, the most important parameter to achieve this is: the melodic contour.

As always, there is a reserve to these sound design recommendations. A sonification can only be successful if it remains a homogeneous set of sounds. Therefore designing distinctly recallable melodies will only be successful if these melodies are sufficiently related to one another to keep the sonification homogeneous.

3.6 Non-Speech Sound at the Human-Computer Interface

The main argument in favour of the use of sound in interactive systems pertains to the way sound works for us in everyday life. For ecological reasons, our auditory system allows us to gather a multiplicity of information from our environment. An essential property of the auditory system is that it allows us to hear what we could not see. Hearing is multidirectional whereas vision is unidirectional. One would hear the sound of a snake first and then try and locate it visually. This phenomenon suggests that sound is a very efficient medium to alert users of the occurrence of an event. Indeed sounds are widely used to convey alarm messages.

Sounds have been used by computer users since computers existed, because computers are part of their users' environment. A common example is that of the hard drive sound. Many users can tell when a saving task is complete from the noise made by their computer. This example demonstrates that the auditory channel is naturally suited to convey information to users in a user interface. In addition, many of today's mobile devices allow sounds to be recorded and played. Therefore, there is no technical limit to using sounds to support interaction in these devices. Conversy and Beaudoin-Lafon divide non-speech sounds in three categories [CBL94]:

- *Alarms*

  These are signals that have priority over any other piece of information. Their goal is to interrupt the current tasks and to inform the user that something requires his/her immediate attention.

- *State and control messages*

  These sounds convey information regarding an ongoing task.

- *Coded Messages*

  These sounds are used to present numerical data as auditory motifs. These sounds are generally more complex and variable than those in the previous categories.

3.6.1 Sonification

The term "sonification" is commonly used as the general way to refer to the audible display of data. Kramer makes a distinction between data-controlled sound, which he refers to as "sonification" and data samples played back directly as sound. He refers to the latter as "audification" [Kra94]. "Auralisation" is also commonly used and relates to the visible representation of data. Other words and phrases have been and continue to be used, such as *audiolisation, acoustic display, virtual audio worlds* and *virtual acoustic display.* The term *sonification* in its most general and informal meaning will be used in the remainder of this thesis.

3.6.2 Earcons

Earcons are structured non-speech sounds that can be combined, transformed, can inherit other earcons properties, and constitute an auditory language of representation (see [BSG89] for an introduction to earcons, and [BWE95] for more up-to-date information). Earcons are constructed from elementary objects of the language that will be referred to as motifs. Blattner and colleagues define earcons as an auditory analogy to icons. They can be used to convey information regarding objects (such as files or folders on a computer), or operations (like copy or delete). Blattner and colleagues divide icons in three categories, although they admit that icons can be represented on a continuous scale [BSG89]:

- *Representational*

  Representational icons are pictures of familiar objects or actions. The advantage of these is that, in principle, they do not need learnt by the user. Unfortunately, all objects and actions do not have a familiar or obvious pictural representation.

- *Abstract*

  An abstract icon is a combination of geometrical shapes that describes an object or an action.

- *Semi-abstract*

  Semi-abstract icons are a combination of abstract and representational icons.

In addition, Blattner and colleagues consider that an icon can be either a basic pictural object, or a composition of basic pictural elements. They present three composition rules to create compound icons [BSG89]:

- *Combination*

  An icon can be created by combination of different icons.

- *Transformation*

  An icon can be created from a slight transformation of an existing icon.

- *Inheritance*

  Inheritance is a transformation that enables the creation of hierarchical relationships between icons.

Blattner and colleagues define earcons as the auditory equivalent of icons. Therefore they share the properties listed in this section.

In practice, earcons are created from simple building blocks called *motifs* [BSG89]. The main issue with this terminology is that it denies earcons the status of a legitimate auditory language. Indeed, a language needs to be built on *elementary* blocks, commonly referred to as the *atoms* of the language. Earlier in this chapter, music has been presented as a language in which the atoms are notes. Earcons exhibit the attributes of an auditory language but lack the notion of atoms. Motifs are not elementary blocks *per se*, they are elementary blocks with regards to an interaction event. Consider these examples:

1. Suppose that a designer decides to use the sequence of notes [C,G] as the building block for the sonification of a text editor. He/she decides to use it as follows:

   - The basic sequence is played every time the application starts.
   - A transposed sequence [D,A] is played every time a file is opened.
   - The latter sequence is reversed when a file is saved.
   - the compound sequence [C,G,G,C] is played when the application is closed.

   This sonification is clearly built on the motif [C,G]

2. Suppose that a designer decides to sonify a desktop interface as follows:

   - the note C is used as feedback to a file selection.

- the note G is used as feedback to a file 'copy'.
- the sequence [C,G] is used as feedback to a file 'paste'

In the first case it is clear that [C,G] is a motif. In the second case, C and G are the motifs of the sonification while [C,G] is an earcon. To avoid any terminological issues, a standard musical terminology will be preferred to the motif/earcon paradigm.

Brewster and colleagues have conducted a number of experiments with earcons from which they have devised guidelines for the design of earcons [BRK96, Bre97]. Theses guidelines provide insight regarding the parameters of earcons that should or should not be used:

Timbre This is the most important grouping factor for earcons. Use musical instrument timbres with multiple harmonics as this helps perception and can avoid masking. These timbres are more recognisable and differentiable.

Pitch and Register The guideline proposed by Brewster and colleagues is coherent with the psychoacoustics results presented earlier: the categorisation of earcons based on absolute pitch judgements should not be expected unless the pitch interval are very large; relative pitch judgements can be used as a cue

Rhythm, duration and tempo Brewster and colleagues suggest that rhythms should be made as different as possible and that putting different numbers of notes in each earcon has proven very efficient.

Intensity Brewster and colleagues recommend not to use intensity as a cue as this is a factor of annoyance. They suggest keeping earcons in a narrow range of intensity. This argument is valid in the sense that sounds are annoying when they are too loud. It is also annoying not to be able to hear a sound because it is not loud enough. However, the latter only applies if the sound needs to be heard clearly by the user. It is arguable that the volume of an earcon may decrease if the relevance of the earcon is low with regards to the user's current task. This principle is commonly applied in the design of many sonifications as Mynatt's *Audio Aura* illustrates [MBW98].

Earcons, or abstract audio motifs, have clearly demonstrated their communication potential. However they have only been applied for the sonification of simple systems which only require the design of a few earcons. The design parameters listed above do not offer sufficient flexibility to undertake the sonification of more complex systems. In addition, the "maximise the differences" type of approach pleaded in the above guidelines may well lead to clearly distinguishable earcons, it may hamper the homogeneity and therefore aesthetic quality of the sonification as a whole. As a consequence, alternative solutions need to be found in order to tackle the design of complex sonifications with abstract sounds.

3.6.3 Auditory Icons

There is an ambiguity in the terminology used in the relevant literature in that earcons have been defined by Blattner as the auditory equivalent of icons, which is what one would expect auditory icons to be. A lot of recent research refers to earcons as what Blattner called abstract earcons, while auditory icons fall under the category of representative earcons. The same terminological approach will be used in the rest of this thesis.

The difference between earcons and auditory icons goes further than the difference between representational and abstract icons. These types of audio messages are fundamentally different in that they relate to two different approaches of sound perception. Earcons require the user to hear sounds in terms of pitch, volume or timbre *e.g.*, in terms of the low-level psychophysical attributes of sounds. On the other hand, auditory icons are perceived in terms of sources *e.g.*, the physical object or process that created the sound. This approach used by Gaver [Gav93a] to define auditory icons is based on an ecological approach of the perception of sound.

As mentioned earlier, the use of auditory icons is limited by the fact that objects such as menu labels do not necessarily have a familiar or obvious auditory representation. Indeed, what auditory icon could represent a menu called "Settings"? This limitation will prove critical in the choice of sounds made during the course of this thesis.

3.6.4 On the Use of Music

Ever since the audio channel has been used to convey information in human/machine interfaces, various authors have addressed the issue of using music [Alt95]. In particular, Alty and Vickers have demonstrated that music can be used to convey hierarchical information that help code and debug programs [AV97]. Rigas and colleagues has investigated the use of music to support graphical information, for both sighted and visually impaired users [RAL97, RA97]. Blattner and Greenberg write: "Music has a communicative aspect not limited to the absolutes of spoken language. Additionally, the "emotional" responses of music, subjective though they may be, can, if harnessed properly, be of tremendous import to the transmission of non-speech audio information" [BG92].

Smoliar points out that when we need a communication medium that involves more than the exchange of words, music is one of the better known disciplines that communicates powerfully through non-verbal means [Smo94]. He argues that since communication is an act of intelligent behaviour by looking at music, rather than natural language, we can more clearly focus this vision of communication as a behavioural process. Alty also highlights the potential of music as a communication medium: "Music is all-pervasive in life and forms a large part of people's daily lives. It is very memorable and durable. Most people are reasonably familiar with the language of music in their own culture. Once learned, tunes are difficult to forget" [Alt95].

There is a long tradition of communicating through non-speech sound like music: horns and bells in Europe: "Hunting horns are an excellent example of signal type non-speech messages (...). These messages included warnings, cheering on the hounds, calls for aid, fanfares for each animal, and so on" [BG92]. And drums in Africa: "Surely one of the most remarkable methods of communication is the talking drum of central Africa (...). The languages spoken in the areas of central Africa where the talking drums evolved are pitched. There are two tones, high and low, that are used variously with each syllable of a word. The talking drums also have tones, high and low, which imitate the tonal patterns of words".

Blattner and Greenberg suggest as well that non-speech sounds like earcons play the role of chorus in Japanese Noh drama [BG92] (in Noh drama, a chorus is part of a coded language that transmit information about the context of the dramatic situation), though one can argue that this language does not take advantage of the specific meanings of music. Indeed, the effectiveness of earcons relies on the fact that people have to learn the structure of the sounds in which information is contained. One can argue that on the contrary, music transmits information without requiring its structure to be understood. This will be discussed in the next section.

Since we have the technology to create any possible sound, it is possible to take advantage of the universal meaning of music to create rich soundscapes that enhance and intensify our computer interfaces. This challenge is still to be met. Accordingly, Gaver outlines that auditory interfaces have so far drawn very little on the possibilities suggested by music [Gav97]. He proposes a possible explanation in that the control needed for the research on auditory interfaces implies a level of explicit articulation, which the complexity of music resists. He continues: "This situation contrasts with designers of multimedia or games environments, who happily exploit music's potential to create mood without needing to articulate exactly how they are doing so". The basis of our approach is that the richness of music belongs to the richness of its meaning rather than in the richness of its structures.

3.6.5 Non-speech sounds in Telephone-Based Interfaces

The use of non-speech sounds in traditional telephony (first category of interfaces defined in Chapter 2) has been poorly addressed in Human Factors research so far. The most significant guidelines are provided by the Ameritech Services standards [SH93]. These recommend the use non-speech sounds when:

- Users do not speak the same language.
- Speech channels are overloaded.
- A spoken message would compromise the security or safety of a situation.
- A specific point in time is indicated, e.g., a special Repeat Dialing tone tells the user that a previously busy phone number is available to ring.
- A spoken message could mask other messages.
- Simple and low cost information presentation is desired.
- Speech output is unavailable.
- The environment possesses characteristics that would mask or block a speech signal.

Nevertheless these guidelines mainly apply to the design of tones. Tones are the only form of non-speech sounds that have been extensively studied and standardised so far, for obvious technical reasons. The European Telecommunications Standards Institute has proposed guidelines for the use of tones on the basis of human factors studies [Ins92]

3.6.6 Non-speech Sounds and Navigation

One of the main problems in interaction with restricted display devices is navigation (see Chapter 2). Previous research has proven that non-speech sounds can be used to represent hierarchically structured information. Barfield and colleagues have created and evaluated simple hierarchical sets of sounds [BRL91]. The aim of their study was to determine whether using sound to represent depth in a hierarchy would assist users in recalling the level of a particular menu option. The earcons that they created was a simple harpsichord sound with a decreasing pitch. The pitch of the main menu sound was an E5, and the pitch of the deepest level was a B4. The experiment showed that the respondent's performance did not improve with auditory cues. This can be explained by the fact that the pitch differences were too small for the respondents to recall the the earcons accurately.

Brewster and colleagues have carried out a series of evaluations that demonstrated that hierarchical earcons can be used to represent small menu hierarchies [BRK96, Bre97]. In their experiments, a 25 node hierarchical menu was used (see Figure 3.7). The results showed that the respondents could identify their location in the hierarchy by listening to the sounds with a very high accuracy. The recall rates were also proved to vary with training and the sound quality. In addition the type of earcons used proved to affect the recall rates of the sounds. The earcons which were best recalled were *compound* earcons (recall rate of 97%). These earcons are based on a numerical representation of the hierarchy: the first child of the first menu item of the hierarchy is named 1.1, and its second child is called 1.1.2. This way, each node of the hierarchy can be represented by a sequence of numbers. In order to create the compound earcons, a motif is designed for the individual numbers and these motifs are then concatenated to form hierarchical earcons. For example, the node 4.3.1 would be represented by the concatenation of motif 4, motif 3 and motif 1.

Although Brewster and colleagues' research suggests that earcons are an effective way of communicating structured information in sound, a number of issues still need to be addressed as far as using non-speech sounds to support navigation in hierarchical menus:

- The effectiveness of non-speech sounds in large hierarchical menus have not been investigated.
  - The sonification of a large menu hierarchy remains to be tackled.
  - The benefit of using non-speech sounds in such a menu has not been proved

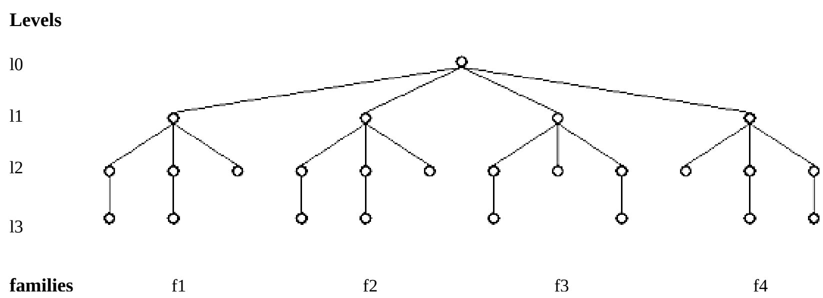


Figure 3.7: Hierarchical structure investigated by Brewster and colleagues [BRK96, Bre97].

- The experiments carried out by Brewster and colleagues [BRK96, Bre97] show that users can locate sounds in a hierarchical structure. This does not prove however that sounds improve the performance of actual navigation tasks.
- The compound earcons, which were best recalled in Brewster's experiment, imply that the user needs to rely on a numerically based mental representation of the hierarchy. The mental process involved in that task suggests that this strategy would not be applicable in a large menu structure. In addition, the duration of this type of sounds increases with the depth of the hierarchy.
- More care needs to be paid to the design of aesthetically pleasing sounds. This is a usability requirement for auditory interfaces.

3.7 Conclusions

This Chapter has introduced the main issues pertaining to the use of non-speech sounds in menu-based interfaces. The designer of auditory interfaces should be aware of these issues and it is hoped that this chapter has presented the necessary information for the reader to proceed to the following chapters.

Chapter 4

An Investigation of Using Music to Provide Navigation Cues

4.1 Introduction

Chapter 3 reviewed ways to use non-speech sounds to improve navigation in hierarchical systems. For instance, earcons (structured non-speech sounds [BG92]) have been shown to be a powerful means to present a hierarchy in sound [BRK96, Bre97]. However, the research by Brewster and colleagues has focused on small hierarchies (25 nodes and four levels including the root's level). Real-world hierarchical menus are usually much larger (the NOKIA *6110* mobile telephone contains several hundreds of nodes and its menu is 6 levels deep). Therefore there is a need to look for additional design parameters to represent menus of this size in sound.

Chapter 3 has also shown that the musical properties of (musical) sounds can be used to convey meaningful information in human-computer interaction. As far as designing auditory navigation cues, research by Brewster and colleagues has shown that *semantic* differences between sounds is the most reliable parameter in the design of non-speech sounds. Indeed, people can easily differentiate musical instruments, as long as they belong to clearly distinct categories of instruments (see Section 3.4). Can *syntactic* differences between non-speech sounds also be used profitably to design a set of hierarchical sounds?

The semantic/syntactic distinction made in the previous paragraph involves a linguistic ap-

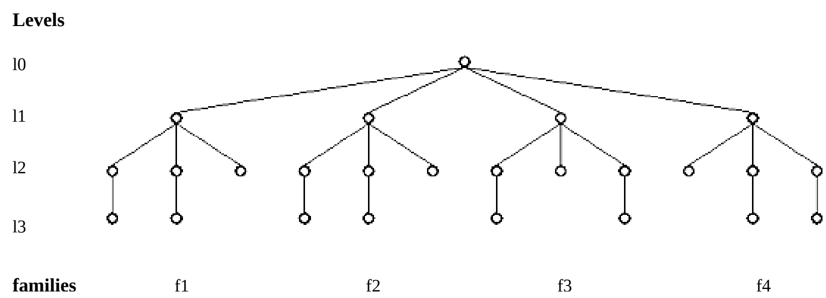


Figure 4.1: Hierarchical structure investigated.

proach to sound design which may need clarification. In a sense, sound design, like music, can be assimilated to devising phrases using an auditory language. The vocabulary, or semantic fields of this language are the various instruments or timbres available to the designer, while the grammar, or syntax, of the language involves all the possible ways to organise the semantic elements of the language. This perspective makes it convenient to refer to the linguistic aspects of music and sound design.

The present chapter describes an experiment that investigates whether people can solely rely on syntactic differences between sounds to remember the structure of a menu hierarchy. A 25 node hierarchy of sounds was created using a single instrument: the piano. The point of the experiment is to show that the hierarchy of sounds could be remembered by the participants, and therefore that syntactic differences between sounds can be used successfully to design hierarchical sets of non-speech sounds. The effect of musical training and experience was also addressed by the experiment.

4.1.1 The Hierarchical Set of Sounds

The 25 node hierarchy was the same as the one used by Brewster et al. [BRK96, Bre97] to allow comparison of the results (see Figure 4.1).

The hierarchy represented in Figure 4.1 is composed of four sub-trees that will be called families. Each of these families contain six nodes and three levels. The root of each family (level one) is represented by a basic motif. The cues for the second level are represented by the level one motif succeeded by a group of notes or chords: for the first node of this level, this group is composed of one note/chord, for the second it is constituted of two notes/chords, and three notes/chords for the third one. For each node of level three in the two first families, the cue is a chord succeeding the sound of the node above it. For each node of level three in families three and four, the cue is a high register melody played in parallel with the sound of the node just above.

Table 4.1 represents the structure of the earcons. The best way to understand the design, however, is to listen to the sounds. These sounds are available on the cd distributed with this thesis (Experiment 1/sonification.html). In Table 4.1, the operator "+" symbolises the concatenation of two sounds, and "/" means that the sounds are played simultaneously. According to this notation, Table 4.1 (last column) shows that families one and two are structured sequentially, as opposed to families three and four whose third level sounds are played in parallel with the corresponding level two sounds. The sounds were intentionally divided in these two categories (sequential and parallel).

As Table 4.1 shows, the design of the four families intended to assign each of the families its own character. The first one is based on a simple minor chord, whereas the second one is based on the very distinguishable scale by tone. The third family was a short rhythmed tune played at a fast tempo whereas the fourth family was based on a II-I-IV cadence.

4.2 The Experiment

The aim of this experiment was to discover if the system of representation investigated would provide results as good as Brewster *et al.*'s [BRK96] experiments for both musicians and non-musicians. Unlike in Brewster and colleagues' experiments, the subjects of the experiment reported here were presented with the 24 sounds of the hierarchy (all of them but the root), as opposed to 12 sounds in [Bre97]. At the end of the experiment, the participants were asked to fill a standard NASA TLX [HS88] workload assessment form (see Section 4.2.2 for more details on the TLX workload test).

The 22 participants were volunteers from the Computing Science Department of the University of Glasgow. They were all research students or members of staff. Eleven of them were classified as being musicians and eleven were classified as being non-musicians. The classification was based on the musical training of the participants and their aptitude at playing a musical instrument. The classification was created for the experiment in the absence of a usable standard classification test. The difficulty of devising a usable (*i.e., short*) test of musical ability has

Family / Level	*one*	*two*	*three*
one	Low register C minor chord	1. + 1 medium register note 2. + 2 medium register, two note chords. The average pitch of the chords increases 3. + 3 medium register, three note chords. The average pitch of the chords increases.	+ 1 high register chord
two	Medium register extract of a scale by tone	1. + 1 high register note 2. + 2 medium register, two note chords. The average pitch of the chords decreases 3. + 3 medium register, three note chords. The average pitch of the chords decreases	+ 1 low register chord
three	Short rhythmed tune	1. + 1 high register chord 2. + 2 high register, two note chords 3. + 3 high register, three note chords	/ 1 high register melody
four	Tune based on a II I IV harmonic progression	1. + 1 medium register note 2. + 2 medium register, two note ascending arpeggio 3. + 3 medium register, three note ascending arpeggio	/ 1 high register melody

Table 4.1: Outline of the global structure of the hierarchical earcons.

been reported by Edwards and colleagues [ECHP00].

4.2.1 Hypotheses

The main hypothesis for this experiment was that the participants should be able to recall the position of a node in the hierarchy by the information contained in the associated sound. If this was correct, high overall recall rates should be observed. The influence of musical ability was also tested. Thus, the experiment was performed on two distinct groups of musicians and non-musicians. Higher recall rates were expected for musicians. In addition, a comparison between sequential and parallel earcons was undertaken. No difference between the sequential and parallel earcons was expected. Finally, the workload survey was carried out for exploratory purposes, therefore no hypothesis was made on these data.

4.2.2 The Experimental Process

The design of the experiment was based on that given in [Bre97] to make the comparison possible. The experiment was divided into a period of training and a period of testing. During the training, participants were presented with the 24 sounds once, and the structure of the hierarchical earcons was exhaustively explained. Then, participants were free to listen to the sounds again as often as they wanted to for five minutes. The test consisted of presenting all the sounds sequentially to the participants in a predefined random order. The twelve first nodes tested were the same as those tested in [Bre97]. The twelve last ones were the remainder. For each node, the participants could listen to the related sound twice. They then selected the node it represented in the hierarchy. At the end of the experiment, each participant was asked to rate the following items of a NASA TLX workload form on a graphical scale: Mental Demand, Physical Demand, Time Pressure, Effort Expended, Performance Level Achieved, Frustration Experienced. In addition, the item "Annoyance Experienced" was added to the questionnaire. This item is often added to the workload questionnaires used by Brewster and colleagues [Bre97] to answer the criticism that sounds are annoying. In the present study however, the annoyance was not related to the presence of the sounds, but to the nature of the tasks.

4.3 Results

The overall recall rate of the sounds was good: 86% of them were correctly recalled. The recall rates for musicians reached 92.8%, and 79.2% for non-musicians. A t-test showed that there was a significant difference between the results of the two groups ($t_{10} = 2.54$, $p = 0.020$). The raw data of this experiment are available in Appendix A.

The recall rate of the first 12 sounds reached 86.3% (92.4% for non-musicians and 81.06% for non-musicians). On these first 12 sounds, the difference between the recall rates of musicians and non-musicians was significant ($t_{10} = 2.51$, $p = 0.021$)[1]. There was no significant difference with the results of the first experiment described in [Bre98] in which the recall rate for the same 12 sounds was 79.9% ($t_{10} = 1.25$, $p = 0.23$). Although the present recall rates are significantly higher than those reported in [Bre97] ($t_{10} = 4.30$, $p = 0.0004$).

There was no significant difference between sequential and parallel earcons since the better recalled family was family one (92.4%), and then came family three (90.15%) then family two (84.1%) and lastly family four (77.3%). These recall rates come in the same order for both groups. The only obvious difference between the results of these groups is in the variance of the recall rates. Indeed, an F-test on the recall rates of both samples showed that non-musicians results were significantly more dispersed ($F_{23} = 0.22$, $p = 0.0003$). Figure 4.2 shows the recall rate of each node of the hierarchy.

The workload survey revealed that the mental demand and effort expanded were rated over 50% (on a scale ranging between 0 and 100) for the overall experiment, for both groups. This rating is just above average, which means that the task was not particularly demanding. Interestingly, all the items of the workload test (except "performance achieved") were rated quite similarly by all the participants, regardless their performance. Some mentioned that the parallel earcons demanded more effort to be recalled, but as the results showed, the recall rates of sequential earcons were not significantly better than the recall rates of parallel earcons.

4.4 Discussion

Several arguments can be put forward to explain these results: The lower results achieved for the fourth family, especially by non-musicians, can be explained by the fact that the basic sound for this family (node 1.4) was by far the longest and the most complex. Consequently, for the lower levels of this family, people could not accurately work out the complex piano stream.

[1] Using a t-test requires the sample distribution to be normal. Throughout this thesis, unless stated otherwise, t-test will only be used when the samples are normally distributed.

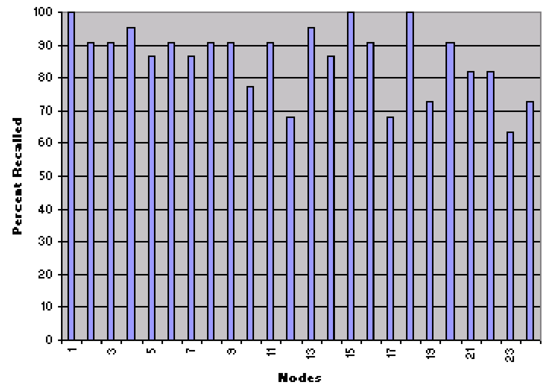


Figure 4.2: Overall recall rates of the 24 nodes of the hierarchy.

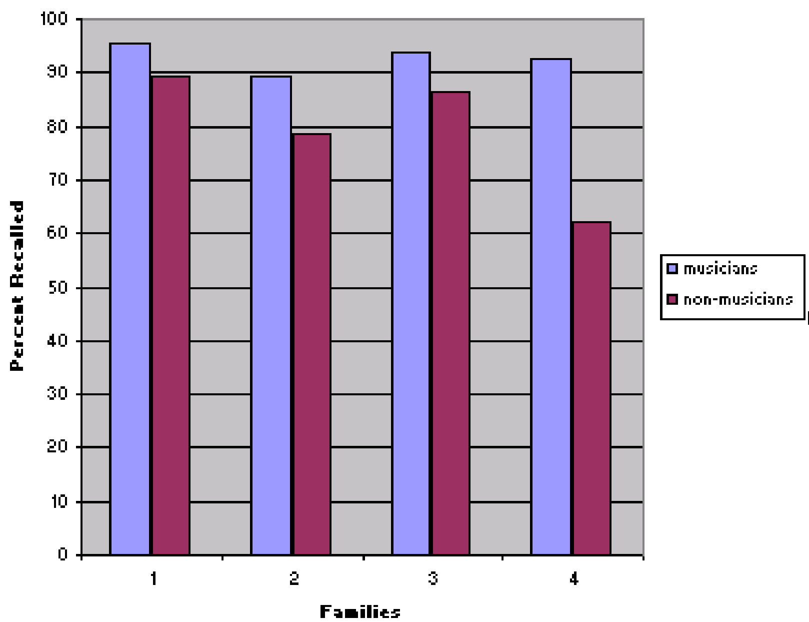


Figure 4.3: Recall rates of the four families of nodes for musicians and non-musicians.

The fact that this basic sound was longer and more complex than the other 1.1, 1.2, and 1.3 sounds could also be the reason why its recall rate was much lower than the other reference sounds recall rates. On the other hand, the defining characteristic of the third family was its fast tempo. The results show that this fast tempo seemed to have helped the participants, maybe by making the sounds shorter. Concerning the difference between the two sequential families, the very distinctively simple chord used as the reference sound for the first family did not give much chance to participants to get confused.

Beyond all the individual differences that occurred in this experiment, it is interesting to look closer at the task participants performed, and derive the arguments that explain their performance. The actual task participants had to perform can roughly be decomposed as follows:

- Remembering the 4 family root sounds
- Understanding the construction rules of the hierarchy

These two tasks involve two different kind of mental activities. Remembering each of the family root sounds requires a correct categorisation of these sounds. The musical meaning of the motive has to be understood. This task was not the easiest to perform by everyone as the motives were all played with the same instrument. The second task involved the capacity to listen analytically to the sounds. When presented with a sound, participants had to extract the relevant information from it. According to the results, two main points arise here:

- On the one hand, people can have difficulties categorising the four root sounds, confusing the different families. This phenomenon can be explained by the fact that these four sounds are highly abstract. In this respect, it is up to people to categorise them. Depending on individual experiences of listening or playing music, they could be very accurately distinguished and then recalled by one, or just sound like piano sounds to another. As suggested in [BRK96], using different instruments for designing sounds should allow us to accentuate distinctiveness of the sounds and so, making their categorisation easier.

- On the other hand, mistakes occur because the extraction of distinct pieces of information from the sounds can be problematic. This was clear for level 4 sounds of the fourth family, the most complex of all. Again, one can presume that using different instruments would facilitate the separation of the relevant streams within each sound.

4.5 Summary

The aim of this experiment was to evaluate new principles to represent hierarchical structures with non-speech audio. The overall recall rate of 86% achieved for a 25 node hierarchy suggest that syntactic features of a musical language of representation could be used as meaningful navigational cues. This results suggest that the earcon design guidelines reviewed in chapter 3.4 can be extended.

However, the search for additional auditory design parameters would be pointless if it did not fit in a clear design framework. At this stage it is unclear whether the results achieved on small hierarchies, without investigating actual navigation tasks, would allow us to tackle more complex real-life menu hierarchies with similar success. Since the objective of this thesis is to improve navigation in restricted display devices, it is time to put the focus on designing sound to support navigation tasks. The research carried out by Brewster and colleagues [BRK96] and that reported in the present chapter only tackle the issue partly. The next chapter will introduce a framework for the design sounds that truely support navigation tasks.

Chapter 5

Supporting Navigation with Non-Speech Sound

5.1 Introduction

So far the thesis has made the two following points: Firstly, navigation is a problematic issue in interfaces with limited graphical display. Secondly, non-speech sounds have the potential to improve navigation in such interfaces. The point of this chapter is to present a framework in which this potential is fulfilled.

Previous attempts at improving navigation with non-speech sound have focused exclusively on the mapping of sounds to the syntactic structure of menus, that is, disregarding the semantic structure of the menu and the actual tasks involved in the navigation process. In contrast, the present chapter presents a novel approach to the design of non-speech sounds based on the very nature of navigation (See Section 2.4 for a review of navigation issues).

5.2 The Design Framework

The review carried out in Section 2.4 pointed out two fundamental features of navigation: On the one hand, navigation is semantically-driven and on the other hand, users use contextual information as navigation cues. These features of navigation are responsible for the problems inherent in navigation in interfaces with restricted display. Therefore the design and implementation of non-speech sounds should aim at tackling these issues. The framework presented in this chapter describes ways to design and implement non-speech sounds that provide semantic and contextual cues lacking in interfaces with restricted display.

The object of the next section will be to present an approach to sound design based on the semantic requirements of navigation in hierarchical structures. The following section will explain how sounds should be designed to provide the additional contextual cues required in the navigation process.

5.3 Supporting Navigation with Auditory Semantic Cues

5.3.1 Visual Analogy

The benefit of auditory semantic cues can be illustrated by a visual analogy. Figure 5.1 represents five menu items grouped in a sample menu. This menu may correspond to the categories related to one's home. The items of the menus relate all relate to various activities, except *car* and *mechanics*, which overlap. Therefore, these items are semantically conflicting. Practically, this means that if a user were to search for a sub-menu called *engine*, they would be as likely to find it in *car* than in *mechanics*.

In this visual example, the ambiguity between *car* and *mechanics* can be resolved with colours. Figure 5.2 illustrates this concept: two very distinct colours (black and white) are allocated to these items. From this point of view, colour becomes a part of the semantic content of the items and help increase their differences when necessary.

Imagine now that a usability study had revealed that the users of the menu had problems categorising the items *e.g.*, the items are perceived as being too similar to each other. In the same visual example one could use shapes to increase the differences between the menu items. Figure 5.3 demonstrates how this can be achieved. In this instance, the shapes increase the differences between the items of the menu and also make *garden* stand out from the rest of the menu.

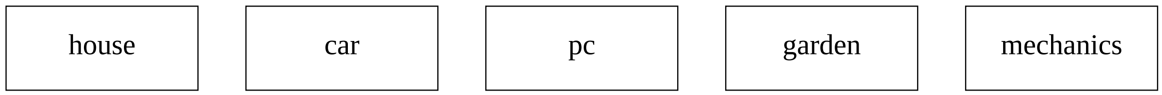


Figure 5.1: Basic set of menu items. the items differ from one another only from the semantic field associated to their label.

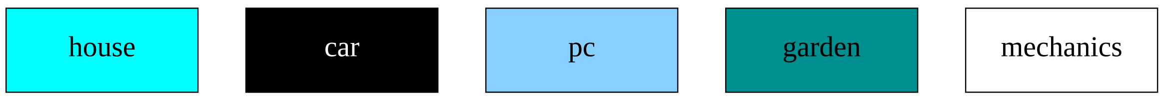


Figure 5.2: Colours are used to overcome the conflict between *car* and *mechanics*. Two very distinct colours (black and white) have been allocated to these two items.

5.3.2 Validity of the Analogy between Visual and Auditory Menus

One may ask whether the visual design carried out in this simple example is analogous to what can be done in auditory design. There are obvious analogies between the visual world and the auditory world, and what has been written above for a visual menu stands almost literally in the auditory world: The semantically conflicting menu items *car* and *mechanics* can be associated to menu items of very distinct colours, and the perceptual differences between all the items can be increased by increasing the differences between the shapes of the sounds allocated to each item. Practically, the colour of a sound relates to its timbre or, in a first approximation, to the instrument that generates it[1]. By the same token, one can associate the shape of a sound to the syntactic structure of this sound. From this point of view, the shape of a sound is related to the shape of its representation on a musical score sheet.

5.3.3 Earcons versus Auditory Icons

It is arguable that the use of auditory semantic cues described above is similar to that of icons. For example, the NOKIA *6110* mobile phone uses icons at the top level of its menu hierarchy. Each item of the top level menu is displayed with an icon *i.e.*, associated to the main menu items (see Figure 5.4). Intuitively, auditory icons stand out as the natural analogy to the non-abstract main menu icons used in the NOKIA *6110*. Nevertheless auditory icons have a number of drawbacks which forbid their use in a large mobile phone menu (see Chapter 3).

- Natural mappings between a sound and a label is rarely possible. Indeed, what auditory

[1] In music, the term colour is used to describe sounds in a more subtle and/or poetical way. The reader interested in the complex issue of sound colour will find an essay by Wayne Slawson interesting [Sla85]

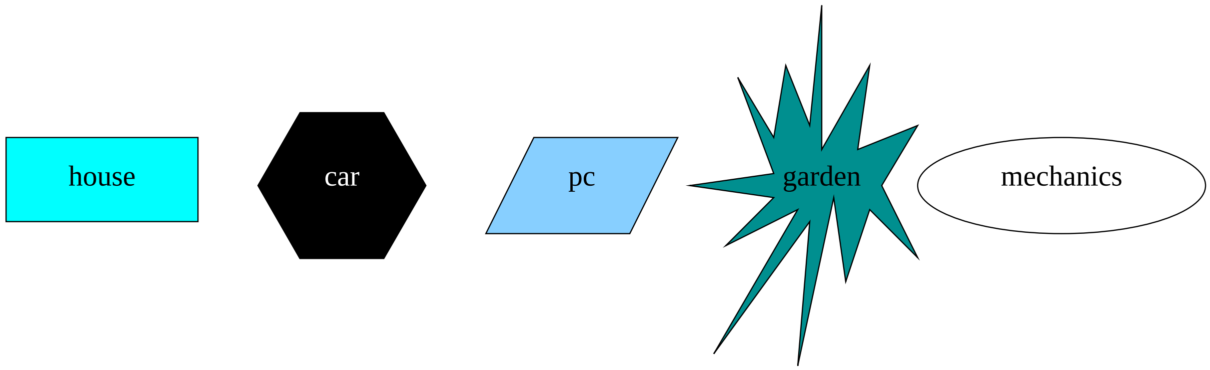


Figure 5.3: Shapes are used to overcome the semantic weaknesses of the menu items. In this example, the choice of shapes makes the item *garden* stand out most.

icon could be used to represent menus such as *Settings*, *Profiles* or *Call Divert*?

- It is difficult (yet not impossible) to create a hierarchical set of sounds of which the root would be an auditory icon. Indeed it is difficult to abstract parameters from representational sounds taken from our environment.

- When a user navigates in a sonified menu, streams of sounds corresponding to the paths he/she takes are implicitly created. The streams generated by a sequence of auditory icons is unlikely to be homogeneous. The lack of flexibility in the design of auditory icons prevents homogeneous soundscapes to be generated.

Therefore the flexibility of abstract sounds such as earcons is needed to carry out the design of a full homogeneous menu sonification. These abstract sounds will be used in the rest of this thesis.

5.3.4 Principle One

The visual menu example presented above demonstrates that non-speech sounds can be used to fulfill two aims regarding the semantic structure of the menu: firstly, to resolve the ambiguity between conflicting menu labels (such as *car* and *mechanics* in Figure 5.2); secondly, to increase the semantic differences between the menu items (see analogy in Figure 5.3). Therefore the first design principle aims at emphasising the categorical character of menu items.

Figure 5.4: Example of a main menu item (*Settings*) displayed on a simulation of the NOKIA *7110* mobile phone. The menu is constituted of a label and an animated icon. On the NOKIA *6110*, the same menu is represented by a label and a static icon.

Principle 1
Use non-speech sounds to increase the perceptive differences between fundamental menu items.

Definition 1
The sounds designed to implement Principle 1 e.g., that increase the perceptive difference between menu items, are referred to as categorical sounds.

5.3.5 Conditions of Use

Principle 1 should be applied especially in large menu structures. This condition of use will be recurrent in this chapter since small hierarchical menus are less likely to cause navigation problems, therefore auditory navigation aids are not required in this context. The question that arises at this point regards the menu items to which the principle applies: Which menu items can be qualified of being *fundamental*? There is no absolute limitation as for the maximum number of sound categories to use, but restricting this number to that of the main menus is recommended. In the NOKIA *6110* mobile phone, for example, such a choice would be consistent with the use of icons in the menu. Indeed, in this phone icons are only used in the main menu.

Categorical sounds can also be used at a deeper level in the menu hierarchy if a menu contains large sub-menus which themselves contain large distinct categories of information. This point is illustrated in Figure 5.5. In this figure, the circled options are the options from the *Security Settings* menu, which is a sub-menu of the *Settings* menu. Such menus may benefit from being allocated categorical sounds because despite their depth in the menu, they are the root of branches that contain very distinct, non-trivial branches. Indeed, the *Call barring* menu contains a significant amount of information pertaining to a very distinct category of tasks from the options contained in the *Closed user group* menu. Conversely, the *Profiles* menu contains options that are semantically similar (see Figure 5.6). All the children of this menu represent an instance of the concept of profile and moreover, their syntactic structure is identical. Consequently these options should not be allocated categorical sounds.

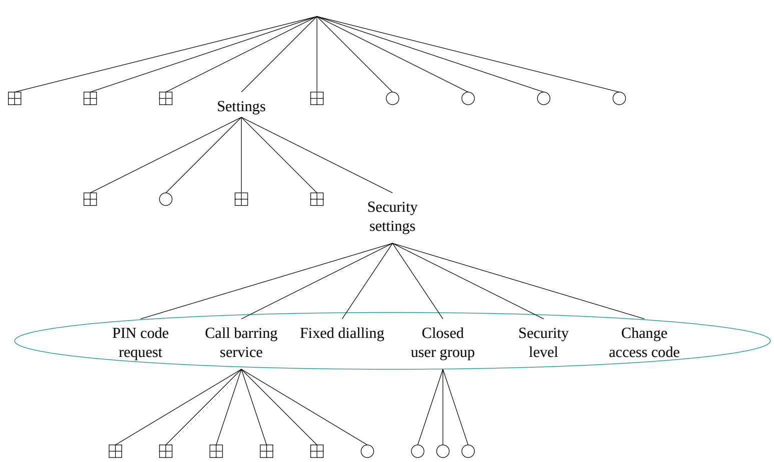


Figure 5.5: The circled options are the options of the *Security Settings* menu of the NOKIA *6110*. This menu contains semantically distinct options that are themselves sub-menus of significant size.

The distinction made between the *Security Settings* menu in Figure 5.5 and the *Profiles* menu

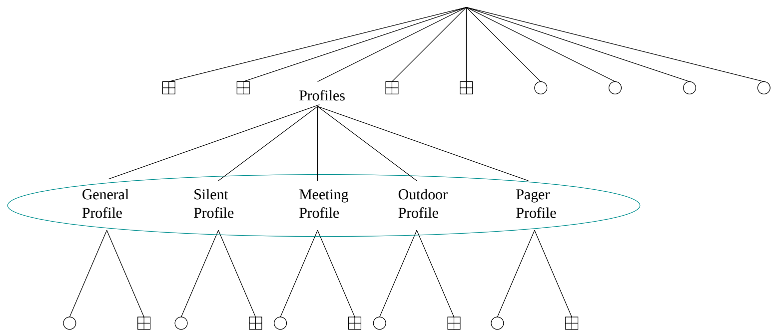


Figure 5.6: The circled options, in the *Profiles* menu of the NOKIA *6110* are syntactically similar. These should not be allocated categorical sounds.

in Figure 5.6 with regards to the semantic links between sub-menus is questionable. In the case of the *Profiles* menu, all the sub-menu can be regarded as instances of *Profiles*, this character being reinforced by the fact that the term “Profile” is present in each of the sub-menu’s labels. From a different angle, the same sub-menus can be regarded as unrelated, since they refer to very distinct interaction contexts; For example *Meeting Profile* and *Outdoor Profile* refer to two opposite behaviours of the interface in terms of the volume of the auditory output of the device. Therefore, a set of menus can represent at the same time, an instance of the same concept, and very distinct categories of actions. If the title of the present section suggests that the distinction should be made on the basis of the semantic relations between menus, no claim is made that prevents any other classification. As Chapter 2 suggested, involving users in the menu item categorisation process may be the best option in this instance.

The structure of the sub-menus involved can also be a decisive factor in choosing between one or the other classification described above. In the case of the *Profiles menu* examined earlier, the structures of all the profile branches are identical. This particularity tends to favour a categorisation of these sub-menus as similar instances of *Profiles* rather than as distinct categories. On the contrary, the syntactic differences between the sub-menus of *Settings* tend to favour a categorical approach of the sub-menus.

5.3.6 Implementation

As far as implementing Principle 1 is concerned, previous research by Brewster and colleagues [Bre97] indicates that the most effective parameter that can be used to distinguish sounds is timbre. In particular, the most effective way to design short distinct sounds is to generate them with distinct instruments. This point is illustrated by Sample 4.

Sample 4

Four syntactically identical motifs played with different instruments: piano, marimba, pizzicato violin and Chatty Aahs[2]

Using distinct instruments to create distinct timbres might sound like a simple practical solution, but it can be a source of errors. For example, an oboe can easily be confused with a clarinet. The solution may be to recommend using instruments belonging to distinct families: for example a violin (string) and a piano (percussion). However such a recommendation would not be accurate; even instruments that belong to distinct families of instruments can easily be confused: for example the sound of a bassoon is not very different from that of a cello, even though these two instruments belong respectively to the families of wind and bowed strings. This problem is even more constraining when low-quality synthesisers are involved. Consequently it is important to choose instruments not according to their name or family, but their psychophysical properties. According to Chapter 3 the fundamental parameters involved in the perception of timbre are:

- Attack.
- Spectrum.
- Spectro-temporal envelope.

Audio sample 5 gives the example of four sounds, syntactically identical, but with strong differences between at least one of the parameters listed above.

Sample 5

Three (non-acoustic) instruments, distinct because of strong differences between at least one of the following parameter: attack, spectrum and spectro-temporal envelope.

The recommendations made so far with regards to the application of Principle 1 suggest that timbre is the most appropriate parameter to manipulate to create categorical sounds. This means that the semantic level of the musical language (timbre) is used at the semantic level of the menu structure. Additionally, it has been shown earlier that, according to psychoacoustics, the semantics of the musical language can be controlled by three main parameters (attack, spectrum and spectro-temporal envelope). This offers a simple practical means of categorising sounds.

[2]This instrument is one of the Ensoniq mr rack instruments.

Nevertheless this categorisation method is not the only valid one. Timbre, which is the heart of the semantic level of the language of music, has been the centre of many composers' preoccupations for a long time. Since orchestration started to become a serious issue, composers began to think of the way to use combinations of instruments to give the resulting sound a desired *form*. The work of composers like Varèse in the twentieth century represented a turning point in the study of sound, and the electronic revolution in music which took place in the 1950's made possible what Varèse was working at: sculpting the sounds with imagination as the only constraint [Viv73]. The work of Schaeffer subscribes to this tradition of sound investigation and *Traité des objets musicaux*[3], Schaeffer's most significant publication, is an insightful study of the nature of auditory and musical objects. He proposes various classifications of sounds as *sonic objects*[4] and *musical objects*[5] [Sch66]. However, Slawson claims that the universality of the classifications proposed by Schaeffer should not be taken for granted [Sla85]. Indeed, Schaeffer's approach is biased by his personal artistic goals and sometimes ignores the psychoacoustical facts of the time.

One can argue that what makes Schaeffer's classification interesting is its subjectivity. Schaeffer uses an analogy with an attic to introduce his classification method: The problem with tidying up an attic is that it involves the classification of heterogeneous objects. Where should an old shirt sit in the attic compared to a bunch of wooden boards and a few bottles? These objects could be universally classified according to their size. This is what psychoacoustics allows one to do. Nonetheless such a classification might not be the most appropriate. Schaeffer suggests that a classification by purpose is sometimes a better solution[6]. His classification clearly aims at electroacoustic composers rather than proposes a universal representation of the auditory world. This approach should be that of the sound designer looking for appropriate sounds; Psychoacoustics provides us with means to classify simple timbres; Gaver's ecological approach has led to a distinct but still meaningful classification of sounds regarding the properties of their source [Gav93b, Gav93a]; the designer may use one of them or devise her own if the context requires it.

Schaeffer's work and that of electronic and electroacoustic composers of the second half of the twentieth century have contributed to the broadening of the perspectives of today's sound

[3] Unfortunately, this book has never been translated in English
[4] *Objets sonores* in the text.
[5] *Objets musicaux* in the text.
[6] *Classer pour l'emploi* in the text.

designers. The more recent work of Trevor Wishart provides further insight on the issue of *sound composition* [Wis94]. In the very specific context of the sonification of a hierarchical menu however, the classification of timbres devised by psychoacoustics should provide the designer with sufficient resources to undertake the design of semantically distinct sounds. This point will be demonstrated in Chapter 7.

It is arguable that operating at the semantic level of the auditory language is not the only means to design categorical sounds. Indeed Chapter 4 has shown that syntactic differences between sounds could be a successful alternative. In fact, musical syntax arguably offers far more possibilities than there are at the semantic level. However, the whole scope of these possibilities will not be explored in details in the present thesis, because these are not the focus of this research. The constraints imposed on the design (temporal in particular: sounds have to be short) imply that only simple musical structures can be employed. These structures can be divided in three categories and, they will be called *motifs* in the remainder of this thesis:

- *Chords*
  A chord is a group of two or more notes played simultaneously.

- *Arpeggios*
  An arpeggio is a group of two or more notes played sequentially.

- *Sequences and superpositions of the above*
  Any motif can be created from chords and arpeggios played simultaneously or sequentially.

An interesting instance of motifs is the *Leitmotif*. Historically, the term leitmotif (German: "leading motive") was first used by writers analysing Richard Wagner's music dramas to characterise recurring themes appearing throughout his pieces. Leitmotives have two dramatic functions: allusion (to dramatic events), and transformation (progression of the drama). Nowadays, the meaning of leitmotif has expanded to themes which carry a dramatic meaning and event to well-known themes. For example, the first four notes of Beethoven's Fifth Symphony can be called a leitmotif. Therefore, one should make a distinction between cultural leitmotifs illustrated by the latter example, and original leitmotifs created to punctuate a dramatic flow.

As Alty suggests, the properties of leitmotifs can be used in auditory interface design [Alt95]. Instead of serving a dramatic purpose, leitmotifs can serve an informative purpose. The first category of leitmotifs (cultural) has the advantage of being easily identifiable and very memorable, but are most of the time overloaded with meaning which does not serve the interface's informational purposes. Moreover they very often carry a strong emotional content which can be detrimental to the interface. The second category of leitmotifs need to be learnt by users, but as they are not loaded with any initial cultural connotation, they are learnt and associated with the information structure only. This approach shows that the first category of informational leitmotifs can be used in a similar way to that of auditory icons whereas the second category is related to earcons. Audio Sample 6 gives an example of two leitmotifs used to sonify two menu items. The first item is called *cds* and the second, *computer*. The first one is associated to the first four notes of Beethoven's Fifth Symphony, whereas the second one is associated to the Intel auditory logo. In this example, the two motifs play the same role, in that they refer the user to the ideas of music and computers respectively[7].

Sample 6

This Sample contains two sounds played sequentially. The first sound is made of the first four notes of Beethoven's Fifth Symphony. This sound is assumed to be associated to an imaginary cd menu. It is important to notice that the exact orchestration is not used in the sample. The syntactic structure of the theme is sufficient to produce the appropriate effect. The second sound of the sequence is the Intel auditory logo. Again, the instruments which plays the motif is not critical, as the theme is well-known enough to be recognised played with any instrument. This sound is assumed to be associated to a menu called computer.

This section has introduced the first principle that addresses the issue of designing sounds to support navigation in interfaces with restricted display. Questions related to the application and implementation of this principle have also been addressed. More details on the implementation of this principle and the following ones will be given in the next two chapters.

5.4 Supporting Navigation with Auditory Contextual Cues

Section 2.4 has shown that contextual information provides strong navigational cues, and that such cues are lacking in a visually-limited display. Figures 5.7 represents the information

[7]It is interesting to notice that the use of sonic logo in branding gives a perfect illustration of the use of the concept of Leitmotif. In advertising, sonic logos are used in a *dramatic* fashion. At the beginning of an advertising drama (*e.g.*, in the first advertising campaign of the brand), the logo is presented in its original form. Then, as the drama unfolds (*e.g.*, campaign after campaign) the logo is transformed to fit in its context, but always reminds the audience (the consumer) the "leading motive" of the drama (the brand!).

available in a standard desktop menu and which are lacking in mobile phone menus. The principles stated in this section aim at compensating for the lack of the following cues:

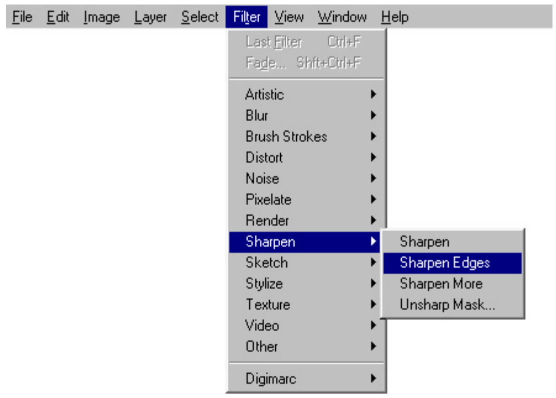


Figure 5.7: Example of standard desktop menu.

- *Depth of current menu item in the hierarchy*

  In a standard desktop menu, it is easy to keep track of one's depth in the hierarchy, because the menu displays all the ancestor menus (see Figure 5.7). In mobile phone menus such as that of the NOKIA *6110*, this kind of persistent information does not exist. Figure 5.8 and 5.9 show that the main way to notify level changes in the NOKIA *6110* relies on a numerical representation of the current menu item, but this cue is far from being available in all mobile phone menus. In addition, in the NOKIA *6110* menu a change of level can be identified from the fact that the style in which the menu item is displayed changes. Again this is not necessarily the case for all devices, and even in the NOKIA *6110*, this cue is not consistently available: the menu items are displayed differently between the main menu and their child menus. Otherwise, the NOKIA *6110* menu items are either displayed one by one on the screen (Figure 5.8) or three by three (Figure 5.9), in a way which is not consistent with the depth of the menus. As seen in Chapter 2, Rosson has demonstrated that level changes were a source of confusion while navigating in a TBI menu [Ros85].

- *Relative position of current item in a list*

  Desktop menus also display the complete list of menus inherited from the father of the current menu (see Figure 5.7), this supplies information to the user concerning the size of the list. Such information, in a device with a graphically-limited display, would help users notice when they reach the end of a list and generally build a better mental model of the menu structure. Some of today's mobile devices, such as the NOKIA *6110*, address this problem by integrating a slider that represents the relative position of the current menu within the list of menu items it belongs to (See Figure 5.8 and 5.9). This feature, however, is not sufficient to provide cues similar to those available in desktop menus. A common interaction problem, illustrating the fact that the slider does not provide sufficient cues, is that of looping. A number of devices (the NOKIA *6110* is one of them) allow users to loop over menus *e.g.*, once they reach the end of a menu list, the next item becomes the first of the list. It is not uncommon that while scanning through a list of menu items, users loop over the list inadvertently. This can easily be explained by the fact that, while scanning, users focus on the menu labels rather than on the slider bar, therefore they do not notice the change of state of this bar. In this instance, auditory cues would solve the problem.

Figure 5.8: Example of a NOKIA *6110* menu in which only one option appears on the screen.

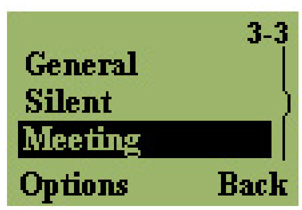


Figure 5.9: Example of a NOKIA *6110* menu in which three options are displayed on one screen. The current menu is highlighted in black, and appears along with two other menus.

Non-speech sound can be used to compensate for the lack of persistent contextual cues described above. Principles 2 and 3 are formulated in this section to address this point.

5.4.1 Principle Two

Principle 2

Use non-speech sounds to supply information regarding changes of level in the menu.

5.4.2 Conditions of Use

As opposed to the first principle, the second principle can be applied in menu structures of virtually unlimited size. It is actually not the the number of nodes in the hierarchy that matters in the present case, but the depth of the menu. The deeper the hierarchy, the greater the need for level change feedback. Therefore, it is recommended to apply this principle when the menu hierarchy is deep.

5.4.3 Implementation Recommendations

As for Principle 1, there are several possible valid approaches to the implementation of Principle 2. On the one hand, one could envisage an implementation based on previous menu sonifications by Brewster and colleagues (from [BRK96]. In the design carried out by Brewster and colleagues, the sound of the root of the hierarchy is the most simple, and the children of a node inherit the properties of their parent and are allocated an additional one. As a consequence, the sounds become more and more complex (and their duration is likely to increase) as the depth of the menu increases. This approach has also been used in the first experimental study reported in this thesis (See Chapter 4). Figure 5.10 shows the menu for which this approach has been followed.

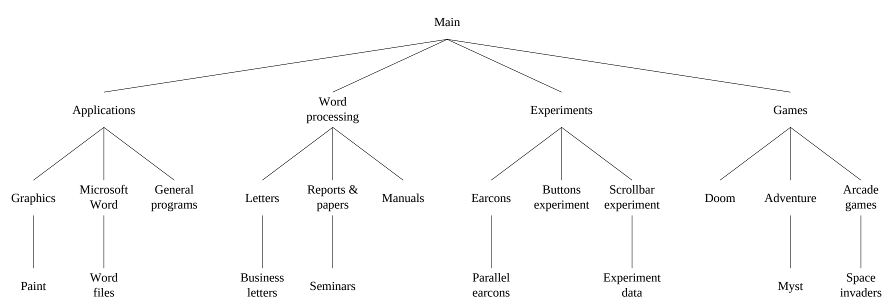


Figure 5.10: The 25 node menu hierarchy sonified by Brewster and colleagues [BRK96].

On the other hand Principle 2 can be implemented by designing less and less complex sounds as the depth of the menu increases. There are two main reasons why one should prefer the latter approach to the former by Brewster and colleagues.

Firstly, auditory streams are created by the movements of users navigating in a menu. If the hierarchical sounds were designed according to the first approach, the *density* of the streams would become too large for the menu to become usable (especially for the deepest nodes of the hierarchy).

Secondly, even for a 25 node hierarchy, Brewster *et al.* have reported that it is difficult to find design parameters for menu items located at the fourth level of the hierarchy. Therefore it is difficult to imagine following this approach to a hierarchy such as that of the NOKIA *6110* (several hundreds of nodes). Even if it were possible, the streams of sounds generated while navigating at the bottom of the menu would be incredibly dense, and unlikely to be usable. Keeping the density and duration of a sonification small is a critical issue because the overall density of a menu sonification is a factor of annoyance (see Samples 7 and 8).

These two points suggest that only the second approach is tenable for the representation of sequences of menu items in sound. Again, a graphical analogy will help making this point clear. Figure 5.11 represents a *balanced* menu sonification whereas Figure 5.12 represents an *unbalanced* sonification. On these figures, the size of a node represents the density of the sound associated to that node. Consequently, the overlap of two nodes illustrates the problem that occurs when dense sounds are played in sequence. It is clear that overlapping is best avoided in Figure 5.11. Let us illustrate, in sound this time, the meaning of overlapping nodes in a menu: Sample 7 represents a stream of overlapping sounds, whereas Sample 8 represents a stream of non-overlapping sounds.

Sample 7

Stream of overlapping sounds (taken from the sonification implemented in Chapter 7).

Sample 8

Stream of non-overlapping sounds (taken from the sonification implemented in Chapter 7).

Sonic density is an obvious parameter to use if one wishes to represent level differences in a menu. Using sounds of variable density would enable Figure 5.11 to be easily implemented in sound. Even though there is no standard definition for the concept of sonic density, it is easy

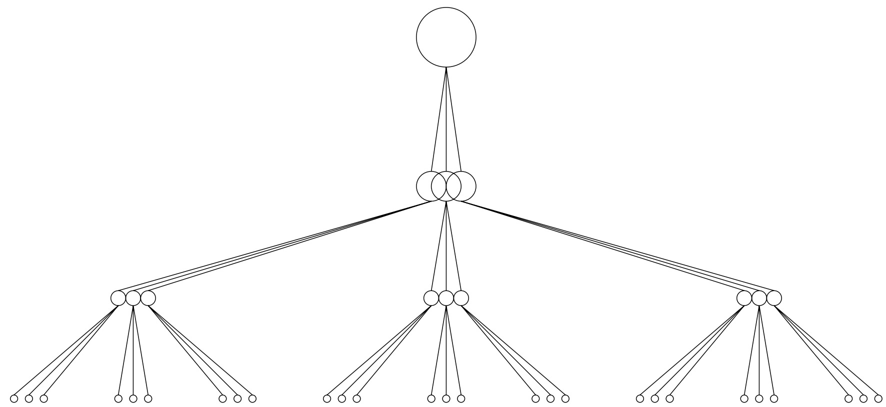


Figure 5.11: Representation of a balanced sonification.

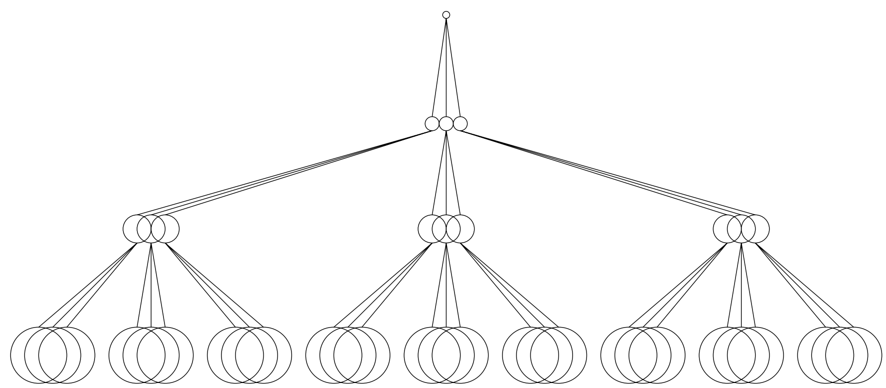


Figure 5.12: Graphical representation of an unbalanced sonification. The overlapping menus at the bottom of the figure illustrate the problem that occurs in audio by browsing a "heavily" sonified menu.

to grasp with an example. Sample 9 illustrates this concept. Density is a multidimensional parameter that involves duration, volume, tempo, note density, harmonicity. Among these parameters, duration should be treated with care, as time plays a critical role in the interaction; the user should not have to wait for a sound to terminate in order to carry on with his/her navigation task.

Sample 9

This sample contains a sequence of four sounds with a decreasing density. The factors that cause the sonic density of the sequence to decrease are the number of instruments, the number of notes, the harmonicity and the volume of the sounds.

If density is not sufficient to convey information about level changes in the menu, or if it is used for a different purpose in the sonification, a different type of auditory cue can be used. An alternative to the use of density is the use of brief percussive sounds. Sample 10 illustrates how four distinct percussions can be chosen to represent four different levels in a menu hierarchy. In Sample 11, a progression based on a single percussion timbre is used to represent the four levels of the same hierarchy.

Sample 10

The sample is a sequence of four percussive sounds: instruments A, B, C and D mapped to four levels of a menu hierarchy.

- *Level 1* $\leftrightarrow$ *instrument A*
- *Level 2* $\leftrightarrow$ *instrument B*
- *Level 3* $\leftrightarrow$ *instrument C*
- *Level 4* $\leftrightarrow$ *instrument D*

Sample 11

This sample is a similar sequence to the previous one in which the four sounds are played by the same percussive sound.

- *Level 1* $\leftrightarrow$ *instrument* A_1
- *Level 2* $\leftrightarrow$ *instrument* A_2
- *Level 3* $\leftrightarrow$ *instrument* A_3
- *Level 4* $\leftrightarrow$ *instrument* A_4

Sample 12 demonstrates how the two previous implementation guidelines for Principle 2 (density and brief percussive sounds) can be combined in a sonification. Sample 12 is taken from the design carried out for the *Call Divert* menu of the NOKIA *6110* reported in Chapter 7.

Sample 12

The present sample contains a sequence of three sounds. The first sound of the sequence is the sound designed for the top node of the hierarchy fragment represented in the figure below (Figure 5.13). The second sound of the sequence corresponds to the middle node, and the third one corresponds to the bottom node. The parameters used to represent level differences are duration and note number e.g., density. They both decrease as the depth increases. A brief percussive sound also provides additional feedback to level changes in the way illustrated by Sample 10.

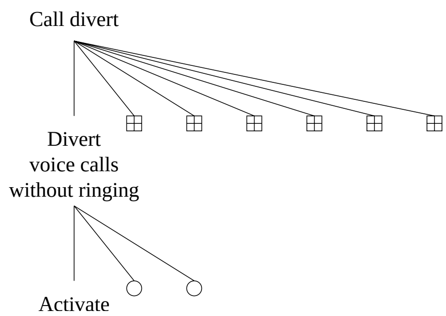


Figure 5.13: Partial structure of the *Call Divert* menu from the NOKIA *6110*.

5.4.4 Principle Three

The introduction of Section 5.4 has shown that mobile phone interfaces do not successfully tell the user the relative position of the current menu options within a menu. This type of information has been shown to be useful to users to encode and recall the location of menus [How94] (this point is illustrated in Figure 5.7). Again, this lack of visual feedback can be compensated for by the use of appropriate auditory cues.

There are two essential characteristics of desktop menus that need to be implemented in the menus of restricted display devices: firstly a cue regarding the position of the current menu option within the menu to which it belongs. Secondly, a cue that indicates the beginning or/and the end of the menu option list. This cue prevents inadvertent looping over the option lists (when the device supports looping over menus). The second point is a particular case of the first one. This issue is the object of the design principle stated below.

Principle 3
Use non-speech sounds to convey information regarding the relative position of the current menu option within a menu.

5.4.5 Conditions of Use

The implementation of Principle 3 depends on one main factor: the size of the branches of which the menu options are the roots *e.g.*, the amount of information contained below the option list. Figure 5.14 represents a menu fragment in which four different types of option lists have been circled. Each of the menus circled in this figure requires a different implementation

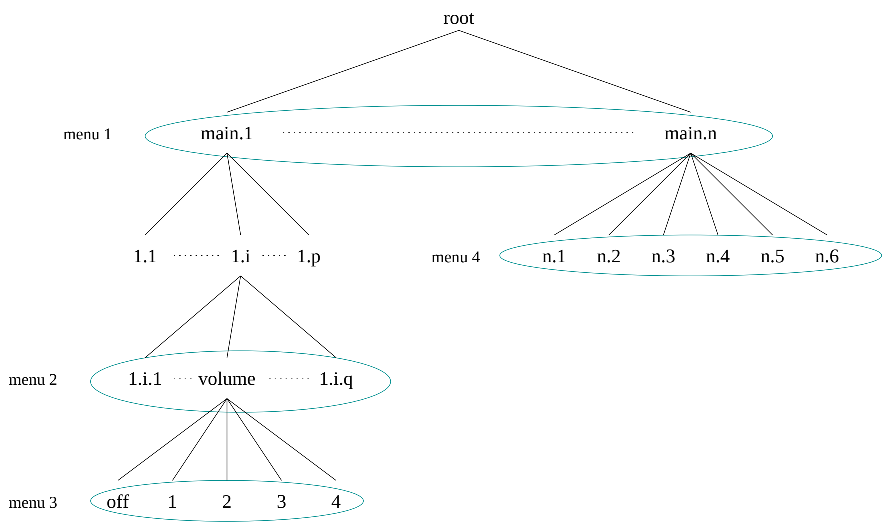


Figure 5.14: Representation of four different types of menus in a menu hierarchy. Four different option lists are circled in this example. 1 - The top list represents the n main options. 2 - The list inherited from main.1 is a list featuring terminal and non terminal options. 3 - The list inherited from *key volume* only involves terminal options. 4 - The list inherited from main.n also only involves terminal options, but at a higher level than the previous list.

of Principle 3. The next section will review four types of implementation which would be appropriate for these different menus.

5.4.6 Implementation Recommendations

The four high-level approaches to the implementation of Principle 3 are the most relevant instances of a design continuum:

1. No sound at all.

   This is a trivial strategy which may be used in a small menu or in a part of a menu that does not require navigational feedback (options located at the bottom of a very deep menu for example). The bottom list (number 3) circled in Figure 5.14 meets these requirements.

2. Same sound for all the menu items.

   This is a trivial non-silent strategy. It can be used when there is no need for navigational cues. However the sounds played while navigating still provide keystroke feedback to the user. This is the sonification style implemented in most of today's mobile phone menus.

Today's mobile phones do not only use a single sound for each list, but they use the same sound for the entire menu (see audio example 13).

Sample 13

This sample represents a user browsing a list of five items of a menu twice. The list being sonified following the second approach (above), the sample is a sequence of ten identical sounds. This menu fragment is represented on Figure 5.15.

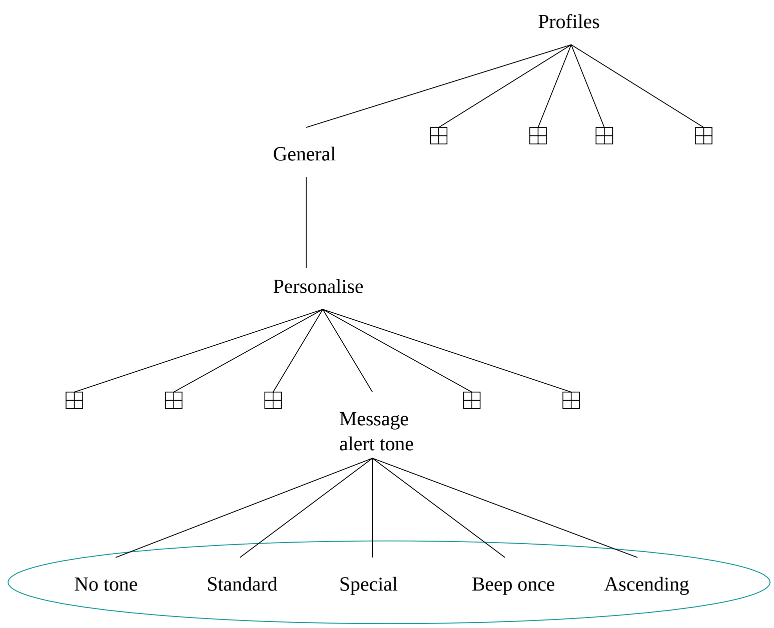


Figure 5.15: Fragment of the NOKIA *6110* menu. The circled menu is located deep within the *Profiles* menu.

3. Two distinct sounds.

   The first sound should be associated to the first option of the list and the second one to the remaining options. This sonification provides the same feedback as the previous one and informs the user when he/she is looping over the list. This is illustrated by Sample 14. The present design strategy could be applied to the top-right circled list of Figure 5.14

   Sample 14

   This sample is a sequence of six sounds that represent a user browsing a 3 option list twice. The sonification has been designed according to the third approach. The relevant menu fragment is represented on Figure 5.16

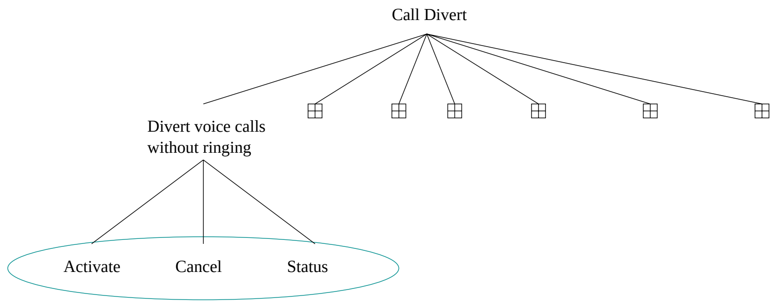


Figure 5.16: Fragment of the NOKIA *6110* menu. The circled menu is located in the *Call Divert* menu.

4. One distinct sound for each option of the list.

   This paradigm can be divided in several sub-cases:

   (a) The sounds follow a progression regardless of the semantic content of the options of the list. This strategy provides the same feedback as the previous one (*i.e.*, notifies the end of a list to prevent unintentional loops); it also carries information regarding the position of the current in the loop 15. Therefore, at this level of sonification only, Principle 3 starts to be fully implemented. This is illustrated in Sample 15:

   Sample 15

   This sample contains a sequence of ten sounds corresponding to the stream produced when a user browses a list of five menu options twice. This example is taken from the design carried out in Chapter 7 for the Profiles menu. The menu fragment is represented in Figure 5.17. The present strategy would be appropriate for the second list from the bottom in Figure 5.14.

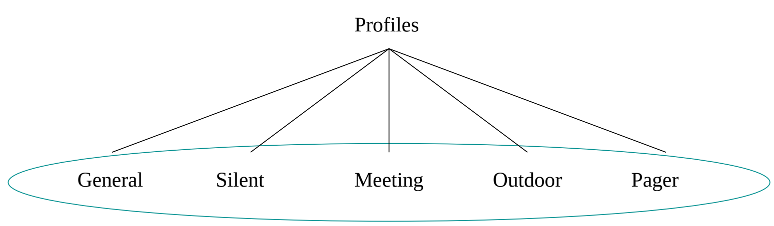


Figure 5.17: Options of the *Profiles* menu available in the NOKIA *6110*.

   (b) The sounds follow a progression similar to the previous one and in addition can be divided in subgroups according to the semantic structure of the menu list. This is illustrated by Sample 16.

Sample 16

This sample contains a sequence of ten sounds produced when a user browses a list of five menu items twice. The stream clearly shows a progression (the average pitch of the motifs increases between the first and the fifth motif). In addition, the formal structure of the motifs reflects the semantic relations between the menu items. Indeed the first two menu items (Alarm Clock and Clock), and the other three items (Call Settings, Phone Settings and Security Settings) are semantically related to each other. This sample is taken from the NOKIA 6110 *sonification evaluated in Chapter 7. Figure 5.18 represents the fragment of menu used in this example.*

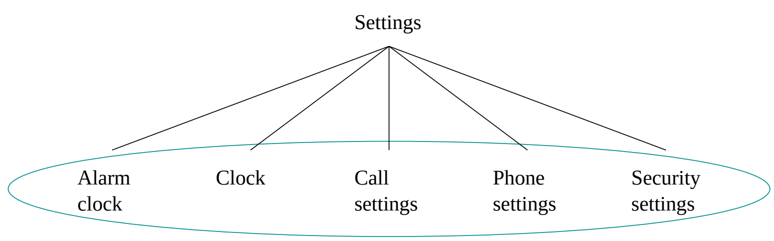


Figure 5.18: Options of the *Settings* menu available in the NOKIA *6110*.

(c) The sounds follow a logical progression providing the feedback required by Principle 3, but more fundamentally support the semantic content of the list. Therefore the sounds associated to the list can be referred to as categorical sounds. Sample 17 illustrates this approach. This last sonification strategy is typically recommended for the top level menu (also referred to as main menu) of a hierarchy. This is the case of the top circled menu of Figure 5.14.

Sample 17

This sample contains a stream of nine motifs generated when a user browses the eight main menu items of the NOKIA 6110 *once and comes back to the first item of the menu. This sonification is also taken from Chapter 7 (more details on th design of this menu will be provided in Chapter 7). Figure 5.19 represents the main menu sonified in this example.*

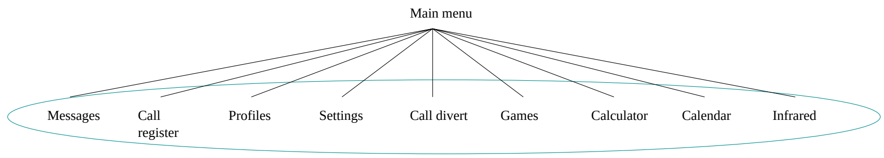


Figure 5.19: NOKIA *6110* main menu.

Again in this instance, the three latter sub-strategies represent three points of a design continuum. At one end of this continuum (3a) the sounds relate to the menu structure from a **syntactic** point of view *i.e.*, regardless of the menu items allocated to the nodes of the list; at the other end of the continuum (3c) the sound follow a less obvious progression but are strongly connected to the **semantic** structure of the menu.

5.4.7 Principle Four

This chapter has so far proposed solutions to compensate for the lack of graphical cues in devices with restricted display by using non-speech sound. The remainder of this section pushes this approach further by proposing additional auditory cues that supply information that standard desktop menus do not provide.

Figure 5.7 shows that desktop menus have the advantage of showing the full path that leads from the root of the menu to the current menu. Nevertheless they do not provide any information about the menus located beneath the current node. Knowing the amount of information available within a menu would be useful while carrying out a search task. This sort of information can not easily be displayed graphically in a desktop menu and certainly not on a mobile telephone where screen space is very limited. In sound, however, it is possible to provide users with a qualitative representation of the amount of sub-menus available beneath the current menu. This is the object of Principle 4.

> **Principle 4**
> **Use non-speech sounds to provide information regarding the size of the branch of which the current menu is the root.**

5.4.8 Conditions of Use

Given the nature of feedback provided by Principle 4, it should be applied to large and heterogeneous menus. Indeed, this information is only needed when the user cannot anticipate the

the size of a branch.

5.4.9 Implementation Recommendations

Sonic density stands as a natural parameter to represent the quantity of information present in a branch. Earlier in the current section it has been proposed that density could be used as a level cue. These two statements are coherent with each other. As density is a multidimensional parameter, the present guideline (using density as an implementation parameter for Principle 4) leaves a large scope of possibilities open to the designer. Sample 9 has showed how density can be used to apply Principle 4.

5.5 Sonification Levels

It has been mentioned earlier in this thesis that the issue of redundancy is critical with regards to the usability of auditory interfaces. The user needs of audio feedback are variable, depending on their experience with the interface. Preferences which are related to aesthetic tastes also play an important role. To address the latter point, various styles of sonifications should be available for users to choose. The former point can be addressed by creating different *levels* of sonifications. This idea is related to the implementation recommendations proposed for Principle 3.

The notion of levels of sonification can be explained with analogy to volume. When sound is present in a system, it is recommended to let users control the volume of the sound. When the volume is too loud for the user's need, it becomes rapidly annoying and conversely, when it is too quiet, it is also annoying. User needs for high or low volumes depend on various factors, in particular the ambient noise. The louder the user's environment is, the louder he/she requires the audio feedback to be. As well as volume, the amount of feedback provided by sound needs to be under the user's control. A novice user may need a lot of feedback to be provided to him/her while and expert user of the same menu may need no auditory navigation aid at all.

5.5.1 Principle Five

Principle 5
Provide the user with several levels of sonification from which to choose from.

5.5.2 Implementation Recommendations

This principle should be applied in any auditory system that meets one or more of the following criteria:

- The system is frequently used by a given user.
- The system is used in different contexts.
- The system involves a large amount of sounds.

As for implementation guidelines, the reader is invited to refer to the description of several levels of sonifications of a mobile phone menu in Chapter 6.

5.6 The Sixth Principle

All the principles presented in this chapter have been created to support the design of sounds that improve navigation in hierarchical menus. These principles often involve maximising the the differences between sounds: for example to help categorise the main options of a menu (Principle 1). However maximising the differences between sounds compromises the homogeneity of the whole sonification. As Chapter 3 has indicated, the lack of homogeneity of a sonification is a factor of annoyance.

As a consequence, the principles stated in the present chapter would miss their target if they were not complemented by an additional recommendation that would almost deserve the status of principle: a *principle of moderation*. The aesthetic qualities of sonifications are often underestimated in the designs carried out in the field of Auditory Display. This section emphasises the necessity to thrive for homogeneity while designing an auditory interface.

5.7 Conclusions

This chapter has proposed a principled approach to the design of sounds to improve navigation in menus. The natures of the menus and tasks involved in the types of interfaces considered in this thesis have been a determining factor in the development of this framework. This theoretical approach provides a novel and necessary basis for the sonification of complex real-life telephone-based menus.

In addition to high level principles, this chapter has also presented a number of implementation guidelines. These guidelines have demonstrated concrete examples of how the principles could be applied. Of course the guidelines can only present a few possible ways to apply the design principles. There are many other ways to implement those principles and the designer's creativity is the only limit to the their utilisation.

Finally, it has been emphasised that the application of these principles should be applied while keeping the overall sonification as homogeneous as possible. The five main principles and the additional principle of "moderation" are summarised below:

1. Use non-speech sounds to increase the perceptive differences between fundamental menu items.
2. Use non-speech sounds to supply information regarding changes of level in the menu.
3. Use non-speech sounds to convey information regarding the relative position of the current menu option within a menu.
4. Use non-speech sounds to provide information regarding the size of the branch of which the current menu is the root.
5. Provide the user with several levels of sonification from which to choose from.
6. Keep the sonification homogeneous while applying the above principles.

Chapter 6

Application I: Design of a Real-World Sonified Mobile Phone

6.1 Introduction

This chapter deals with the application of the principles and guidelines defined in Chapter 5 to the sonification of a mobile phone prototype menu. Chapter 5 has proposed a set of principles and implementation guidelines to support navigation in visually restricted interfaces. The assumptions regarding the interaction constraints in that chapter were dealing with the visual limitations of the device and no assumptions were made concerning its audio capabilities. Such an assumption is justified by the fact that the development of newer generations of mobile devices shows a clear trend towards greater audio capabilities, not only devoted to the reproduction of speech, but also to the presentation of pre-recorded audio samples (MP3 players and radios, for example).

The work described in the present chapter, however, tackles the issue of designing sounds in a technically constrained context. There are a number of reasons why such an implementation has been carried out: firstly, it was the only technology available for the development of a real prototype when this research was undertaken; secondly, despite the improvement of the audio quality of the new generations of mobile devices, it is likely that a number of categories of (low-end) mobile devices will still involve a poor audio quality. Designing for the "worst case scenario" provides valuable insight into the principles investigated in this thesis, and is likely to improve the principles devised so far and make them applicable to a wider range of contexts.

The first part of this chapter is dedicated to the presentation of the technical constraints that had to be dealt with in the design. it is demonstrated that the principles and guidelines proposed in the previous chapter enable the implementation of a sonification, despite the sound design constraints imposed by the audio capability of the device. Even though design is the focus of this chapter rather that evaluation, the final part of this chapter accounts of the evaluation of this prototype carried out by Helle and colleagues [HLML01].

6.2 The Technical Constraints

The development of the prototype described in this chapter was based on the NOKIA *8210* mobile telephone menu. The NOKIA *8210* mobile telephone menu is a large menu comprising several hundred nodes. The complete structure of this menu is presented in Appendix B. In this device, without the use of the headset, sound production is possible thanks to a basic pulse synthesiser driving a buzzer. The properties of this hardware and their implications on the application of the principles and guidelines defined in Chapter 5 are listed below:

- **Single instrument**

  The only sound available is a buzzer tone (buzzer tones are the sounds produced by most mobile phones when a key is pressed. Their wave form is a square wave). The timbre of the tone *i.e.*, the properties of the square shape of the signal, cannot be modified. The fact that only one instrument is available for the sonification is a serious limitation to the design. Indeed one of the main implementation guidelines for Principle 1 suggests that using distinct timbres is the most appropriate solution to design categorical sounds. However, Chapter 4 provides some help in this case as for the design of hierarchical menus using other parameters than timbre.

- **Poor aesthetical quality of the sound**

  Buzzer tones are widely used to provide brief audio feedback in many devices. This fact is not a matter of aesthetic choice but is the direct consequence of technical limitations. With today's electronic devices, users have become accustomed to an increasingly richer variety of more subtle musical instruments than a buzzer tone (sonification sets provided with Microsoft Windows Operating systems for example). Consequently, the shift in the users' expectations regarding the timbres used in a complex sonification is likely to affect their appreciation of the prototype sonification developed here. To summarise, whatever the design, it is likely to be badly rated by users because of the "musical" instrument used to generate it.

- **Poor control over the sound**

  Not only is there just a single instrument available, but the control allowed over it is extremely limited:

  - Limited volume control: five volume levels are available, and the volume cannot be changed within a motif.
  - No control over the envelope.
  - No control over the timbre.
  - No dynamic control. The synthesis parameter of the sounds are determined before the sound is played. Once a sound has started, its parameters (including pitch and volume) cannot be modified.

- **Monophonic**

  This constraint prohibits the usage of chords. This is a serious limitation as it forces all the notes to be played in sequence, and by nature sequences take more time than chords. This implies that only minimal messages can be played in a brief interval of time.

- **Limited pitch range**

  Only a narrow pitch range is available: 440 - 2096 Hz. This prevents the use of the low register. The register available extends on two octaves from A440 to C2093. Unfortunately, buzzer tones tend to be more irritating and even alarming in this range of frequencies.

The list drawn out above suggests that it is difficult to imagine a more hostile context for sound design. However, the next section shows that even with severe constraints, it is possible to apply the principles stated in Chapter 5 and sonify a complex menu hierarchy.

6.3 Overcoming the Technical Constraints

Some of the guidelines formulated in the previous chapter involve the use of different instruments, or the use of chords, for example. Such guidelines are not applicable in the current instance, because only one instrument is available and this instrument is monophonic. However, when an implementation guideline is not applicable, the high level principle it relates to always is.

To find the most appropriate way to apply these principles in the present context, an exploratory investigation of sound design using buzzer tones is useful. In getting familiar with the "sonic world of square waves", a lot can be learned from the sound design work carried out two decades ago, especially in the computer games industry. The audio capability of the game platforms at time were similar to those of today's mobile phones. The aesthetic characteristics of the sound-tracks and sounds designed for games in this period of time are easily identifiable and their authors achieved excellent designs given the constraints they had to deal with (see games such as *space invaders* or other classics implemented on platforms like the Amstrad or Atari micro-computers).

6.3.1 Musical Parameters of the Design

The constraints highlighted in the previous section demonstrate that one can only operate at the syntactic level of the auditory language (*e.g.*, only the formal structure of the sounds can vary, not the instrument that generates them). Therefore, one can only use the following main syntactic parameters in the design:

- **Melody**

  Contour is the most obvious variable to use in this context. First of all, the sound generator being monophonic, brief melodies are the only musical structures one can deal with. According to the review carried out in Chapter 3, differentiating main classes of contours does not require users to possess any special musical ability; in a very limited interval of time, it is not possible to make use of sophisticated melodic structures.

- **Harmony**

  Auditory phrases can be implemented with harmonic progressions. Harmonic differences can also be used to create categorical sounds. Therefore, depending on the progression or differences, harmony can be used as a parameter to implement both Principles 1 and 2.

- **Tonality**

  A change of tonality represents a break in the tonal structure of a sonification. In tonal music, tonality changes appear at carefully chosen times in a piece. Tonality changes should be preferably used to create categorical sounds.

- **Rhythm, Meter and Tempo**

  Chapter 5 has shown that the scope of action regarding the temporal properties of the sonification is very narrow: the duration allowed for a sound cannot exceed the duration that a user spends on the relevant menu item.

6.3.2 Dealing with Monophony

The solution that has been found to overcome the problem of monophony is the use of *pseudo-chords*. Pseudo-chords can be defined as arpeggios played rapidly. Pseudo-chords have the "taste" of chords, but are in fact arpeggios. This effect is similar to that of a violinist playing a chord on more than two strings. Creating these sonic objects is possible because the buzzer that generates the sounds in the prototype can play notes as short as 25ms.

To illustrate this concept, a brief piece of music has been composed. The piece is compound of two simultaneous voices. The original version of this piece is contained in Sample 18. The monophonic version of the same piece is contained in Sample 19. The latter version involves pseudo-chords.

Sample 18

Polyphonic version of a two-voice tune. The first two bars of the tune are represented in Figure 6.1.

Figure 6.1: First two bars of a polyphonic tune.

Sample 19

Monophonic version of the same two-voice tune as above. An almost polyphonic effect is achieved using pseudo-chords. The first two bars of the tune are presented in Figure 6.2.

Figure 6.2: First two bars of a pseudo-polyphonic tune.

The notions of pseudo-chord and *pseudo-polyphony*[1] are related to the fusion/segregation phenomenon described in Chapter 3. This phenomenon has been comprehensively investigated by Bregman [BC71, Bre90]. In Sample 19, the fusion effect is achieved because the tempo at which the tune is played is fast and the pitch difference between the two voices is large. Accordingly, in the middle of the tune, when the pitch difference diminishes, the fusion effect becomes less obvious and the piece starts sounding less like its polyphonic counterpart.

6.3.3 Putting the Emphasis on Musical Movement

In order to overcome the absence of semantic parameters available in the auditory language (single instrument available), the emphasis has to be put on the syntactic elements of the language. The study reported in Chapter 4 shows that people can rely on the syntactic content of a set of sounds designed with only one instrument in order to locate their position in a hierarchical menu. In that study however, polyphonic sounds were used. Constrained by monophony, or pseudo-polyphony, melodic and harmonic progressions become an essential parameter of the design.

6.4 Sonification Levels

According to Principle 5, the user should have the option to choose from a set of sonifications to adapt to his/her needs and tastes. This section presents a range of sonifications (set I to set IV) that provide an increasing amount of feedback to the user.

- **Set I**
  The first set is a trivial sonification: all the sounds are identical. This is the sonification currently used in the NOKIA *8210*. This level of sonification is mentioned to highlight the progression between the level of audio feedback provided by real-world devices, and the sonifications described below.

[1] pseudo-polyphony can be defined as the “almost” polyphonic effect achieved by using pseudo-chords.

- **Set II**

  In the second set, each main menu is represented by one minimal sound e.g., a note. Each main menu sound is then used in each of the main menu's sub-menus. The volume can also be used as a level cue.

- **Set III**

  In this sonification the motifs used are distinct for each main menu, and also across subsequent sub-menus.

- **Set IV**

  This set uses the same paradigm as above. In addition, the sonification provides information about the amount of information contained below the current menu item.

For technical reasons, only one sonification could be implemented in the prototype developed by Nokia. This sonification is the fourth set described above and will be described extensively in the next section. This sonification implements the first four design principles stated in Chapter 5.

6.5 Design of a Set of Sounds for the Prototype

The present section focusses on the application of the principles divised in Chapter 5 for the sonification of a real mobile telephone prototype. The resulting sonification is available on the CD attached to this thesis (See Appendix D.4.1 for instructions on how to listen to the prototype sonification).

6.5.1 Main Menus

In Chapter 5 it was suggested that, in large menu structures, the main menu items should be represented by categorical sounds (which is the case of the NOKIA *8210* menu). In addition, Principle 3, which deals with the auditory feedback for browsing a list of menu items, encourages the use of a progression of sounds in order to sonify a list of items. Besides, the implementation guidelines for that principle suggest that categorical sounds are nothing but a form of progression between sounds with strong semantic differences. Because it is not possible to operate at the semantic level of the design, Section 6.3.3 explains that it is necessary to rely on sound progressions or musical movements rather than on semantic differences between sounds, even for the main menu items.

As a consequence, the main menu is sonified by a progression of syntactically distinct motifs. Here is a description of the various stages of the design that lead to the main menu sonification: The first stage of the design involves choosing a scale on which the progression is based: a pentatonic scale (see Sample 20). The choice of this scale is one of these non-scientific choices that occur in sound design. However, it is not arbitrary. Melodies made of ascending or descending major scales have a simplistic and therefore musically dull connotation. asending and descending minor scales are similar but with a melancolic or sinister aspect. As for chromatic scales, they are probably the least pleasant. Scales composed of fewer notes allow for the design of more interesting motifs. The pentatonic scale presented here constitutes an easy way to design easily identifiable and resonably plesant short asending or descending motifs.

Sample 20

A minor pentatonic scale used as the basis for the musical movement involved in the sonification.

Figure 6.3: Pentatonic scale, transcription of Sample 20.

There are eleven main menu items in the NOKIA *8210*, therefore eleven sounds have to be created. Sample 21 illustrates the first step of the design: the choice of a basic progression of sounds based on the scale contained in Sample 20. Because of the scale in which the notes of this melodic progression have been chosen, the progression carries harmonic pieces of information that are utilised in the next step of the design. This harmonic content is such that the progression would accommodate various rhythmic patterns. Sample 22 provides an example of possible adequate rhythmic pattern.

Sample 21

Melodic progression created to sonify the main menus. The sequence can be divided in two sub-sequences, an asending sequence of note (first five notes) based on the chosen pentatonic scale and a descending sequence based on the same scale (remaining six notes). This division reflects the basic structure of the main menu in which contains, on the one hand, five large branches and, on the other hand, six less significant branches.

Figure 6.4: Transcription of Sample 21.

Sample 22

Example of a rhythm matching the melodic progression of Sample 21.

Figure 6.5: Transcription of Sample 22.

The third step of the design corresponds to the development of the harmonic content of the progression, based on the initial melodic phrase. The chosen implementation of this is represented by Sample 23. At this point of the design, the eleven sounds are divided in only two distinct groups. The first group contains the first five menu items whereas the second group contains the last six items.

The justification for this first categorisation lies in the structure of the menu hierarchy: in first approximation, the main menu items can be divided in two groups. In the first group, the first five menu items are those which are the root of a branch of a significant size. On the other hand, the second group contains items which are the root of much smaller branches. This categorisation is coherent with Principle 4.

Sample 23

Development of the melodic progression of Sample 21 by application of Principle 4.

Figure 6.6: Transcription of Sample 23.

In the third step of the design the main menu motif sequence had to be broken down into more categories, with reference to Principle 1. This has been achieved by transposing the sounds associated to the third and fourth menus. The resulting sound sequence is contained in Sample 24).

Sample 24

The set of sounds presented in this sample is the result of the application of Principle 1. The notes forming the third and fourth motifs have been transposed to achieve an easier distinction with the first two motifs: the first note of these former motifs has been transposed up by two octaves while the other two notes have been transposed up by one octave. This resulted in a modification of the contour of the motifs and consequently increased the difference between the first two motifs and the next two.

Figure 6.7: Transcription of Sample 24.

The fifth step of the design involves the refined application of Principle 4. In this refinement, the main menu is broken down into three categories with regards to the size of the menu main menu items: The last six items are smaller than the first five, therefore the density of their sound is kept small. In opposition to this, the third and fourth items are significantly larger than the other menus. Therefore the density of their motifs is increased. Sample 25 illustrates the modification made to Sample 24 by application of Principle 4.

Sample 25

This Sample represents the modifications made to the previous set of motifs by re-applying Principle 4. This second, finer application of Principle 4 involved increasing the density of the motifs of the two significantly larger branches of the menu: the third and fourth motifs of the sequence. To increase their density without increasing their duration, the notes forming these motifs have been replaced by two-note chords. This was achieved by combining the motifs obtained after the second step of the design with the motifs resulting from the third step of the design.

Figure 6.8: Transcription of Sample 25.

The sound set produced in the last stage of the design contains two polyphonic sounds (for the third and fourth menus). The last stage of the design of the main menu sounds involves modifying these two motifs, to meet the constraint of monophony. Pseudo-chords, which were introduced earlier in this chapter, have been created for this purpose. The transformation of the chords of the last stage into pseudo-chords is illustrated in Sample 26.

Sample 26

Transformation of the third and fourth motifs of the previous sample, to meet the monophony constraint. The chords present in the previous sample have been replaced with pseudo-chords.

Figure 6.9: Transcription of Sample 26.

This section has described in details all the stages involved in the design of the main menu sounds. This set of sounds will be used on the next section as the basis for the sonification of the main menu branches.

6.5.2 Messages Menu

The *Messages* menu is the first of the main menus. It consists of nine sub-menus among which two have sub-menus themselves. The challenge of designing sounds for this branch (and the next one) consists of ensuring that all the design principles stated in Chapter 5 are applied. First of all, Principle 1 does not need to be applied within this branch, according to Section 5.4.2. The cues for level changes recommended by Principle 2 involve a change of density of the sounds: the number of notes and/or their duration decreases as the depth of the hierarchy increases. Principle 3 is applied everywhere in the branch except at the bottom level: except at that level, the lists of items contain at least two different motifs. Consequently the user knows, to a certain extent, what is his/her position within the menu list. Finally, Principle 4 is applied by mapping the number of notes of a motif to the size of the associated menu. For example, Figure 6.10 shows that among the nine motifs of the second level of the menu (the root being at the first level), the two that are made of two notes correspond to the menus that contain sub-menus. The design rules for each level of the branch are summarised below:

- *Level 1*

  This is the level of the root. The design of this motif is detailed in the previous section. The duration of the notes constituting this motif is 96 ms[2].

- *Level 2*

  The motifs are made of one note if they are terminal, two otherwise. They constitute an ascending progression based on the scale represented in Sample 20. The duration of an eighth note at this level is 64 ms.

- *Level 3*

  All the motifs are sonified by a single note motif. The note associated to the first node of a list of third level items is the first note of the parent of the node in the hierarchy, whereas the note associated to the other items of the list is the second note of the parent's motif. For example, using a numerical representation of the menu items, and starting with number for the root, the motif for node 1.5 is a sequence [E,G]. Consequently, the motif for the node 1.5.1 is made of a single note [E] and the motif for nodes 1.5.2 and 1.5.3 is made of a single note [G]. The duration of an eighth note at this level is 48 ms.

- *Level 4*

  The nodes' motifs are all similar to their parents'. The duration of an eighth note at this level is 32 ms.

For consistency's sake, the same approach was followed to design all the main menu branches. Therefore, only the outcome of the design will be presented in the next sections.

6.5.3 Call Register Menu

The approach to the sonification of the *Call Register* menu is consistent with that of the sonification of the messages menu; the description made above for Level 1 and Level 2 stands for *Call Register* as well. This is illustrated by Figure 6.11.

6.5.4 Profiles Menu

The *Profiles* menu and the *Settings* menu are much larger that the two previous ones. However they differ from one another by the fact that the structure of the *Profiles* menu is more regular

[2]The duration of eighth notes is set 96 ms because 96 is the closest multiple of 4 and 6 to 100. The desired eighth note duration being in the range of 100 ms, 96 ms enables 16th and 32nd notes to have integer durations (48 and 24 ms respectively), as well 8th and 16th triplets (64 and 32 ms respectively)

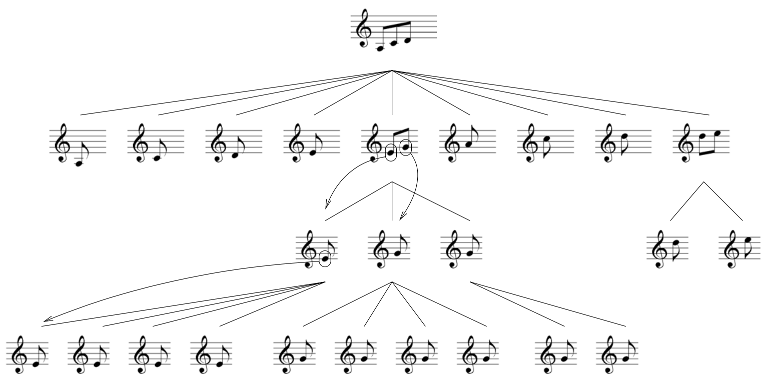


Figure 6.10: Sonification of the *Messages* menu.

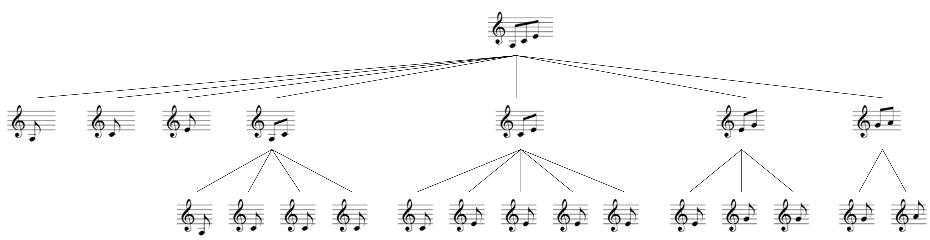


Figure 6.11: Sonification of the *Call Register* menu.

than that of the *settings* menu. Indeed, the *Profiles* menu is made of five identical branches whereas the *settings* menu is made of branches of very distinct structures. Figure 6.12 represents the sonification of the first branch of the *profiles* menu. The sonification of the other branches of this menu have been obtained from a simple transposition: the motifs of the second branch are one degree higher in the pentatonic scale used in this design. The motifs of the third branch are one degree higher than the motifs of the second branch, and so on (See multimedia simulation of the menu available on the CD attached to this thesis, details in Appendix D.4.1).

6.5.5 Settings Menu

The *Settings* menu has a large and irregular hierarchical structure. Figure 6.13 represents the sonification of this branch.

6.5.6 Call Divert Menu

The approach to the design of the *Call Divert* sounds was the same as that used previously for the *Messages* and *Call Register* menu. In Figure 6.14 one can notice that there are eight items

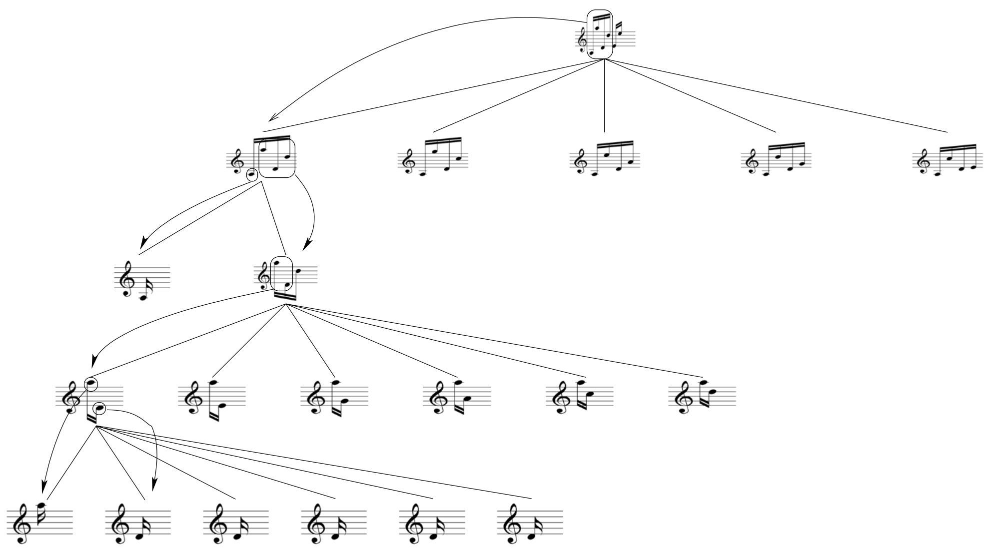


Figure 6.12: Sonification of the *Profiles* menu.

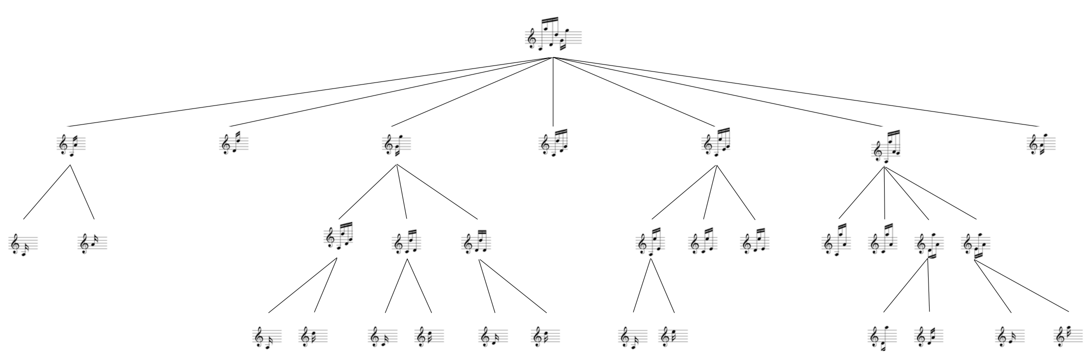


Figure 6.13: Sonification of the *Settings* menu.

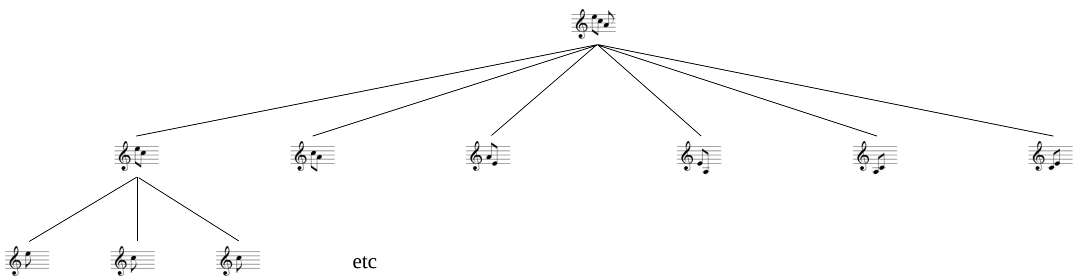


Figure 6.14: Sonification of the *Call Divert* menu.

at the first level of the menu. The first four follow a descending motion based on the Sample 20 scale, whereas the remaining four follow an ascending movement.

6.6 Evaluation

The work described in this chapter could be presented from two different angles: From a purely academic point of view, the design presented here was was carried out to prove the robustness of the design principles defined in Chapter 5; practically, the work was motivated by the fact that the design would be implemented in several real mobile telephone prototypes, the prototypes given to users and some feedback would be collected.

The evaluation of the prototypes has been perfomed by a team of colleagues from Nokia and the results of their evaluation has been published [HLML01]. The outcome of this evaluation, that was not carried out by the author but what is relevant is summarised here (see the study by Helle and colleagues for more information on the experiment).

Seventeen test users, 7 males and 10 females, aged from 24 to 45, all Nokia employees working in office environment in various occupations and experienced mobile phone users, were given a prototype for a period of three weeks. The results of the study indicate that, predictably, the prototype was negatively received due to the type of sounds used for the sonification. Some positive comments however were also made by some users [HLML01]:

'In the user study, it was clearly indicated that the prototype implementation of sounds was not well suitable for regular usage as it was considered too disturbing by most users. A more sophisticated sound reproduction system would be needed, and sound radiation should be limited by some means, or the sounds should be restricted to certain usage situations. However, we received also signs of usefulness from a few test users. Those may encourage developing the concept further."

6.7 Conclusions

The present chapter described the process of designing a sonification for a real mobile telephone prototype. The mobile telephone prototype was based on the NOKIA *8210* mobile telephone, which contains a large hierarchical menu. This study demonstrates that the principles proposed in Chapter 5 can be applied in a very constrained context.

The design constraints faced in this study were linked to the properties of the prototype sound generator: the device buzzer was used to generate the audio content of the menu. The buzzer is the component used to play the ringing tones and alerts generated by the device. The main limitations of this sound generation medium is that sounds can only be played in a narrow band of frequency, and polyphonic sounds cannot be produced.

As a result of these constraints, some of the implementation guidelines could not be applied. Therefore, the design was carried out to meet the requirements of the high-level principles. The design process showed how the constraints can be dealt with to meet the high level requirements. Here is a summary of the techniques used to meet these requirements:

Principle 1 The syntactic differences between the main menu options were emphasised by tonal differences between the sounds allocated to these items. Throughout this design, pseudo-polyphony was used to overcome the constraint of monophony.

Principle 2 Level changes were notified by density changes: the duration and the number of notes of the sounds decreased while the depth of the related node in the hierarchy increased.

Principle 3 Except from the menu lists located at the bottom of the hierarchy, each menu list was represented by a set of sounds that formed a clear progression: for example, the set of sounds could involve a series of notes of increasing pitch.

Principle 4 This principle was implemented in conjunction with Principle 2: differences of densities were used to represent level changes, while sonic density was used to represent the amount of information available in the branches of the hierarchy.

Principle 5 This principle suggests that different levels of sonification should be made available to users. Consequently, four sets of sounds were specified for the prototype.

The application presented in this chapter has illustrated the implementation of the principles of Chapter 5 in a challenging context. In addition, the description of the design constitutes a good basis for the implementation of further sonifications in the same context. Indeed, it is probable that, in the future, some categories of mobile devices will still have similar audio capabilities to those of the prototype investigated in this study and require a similar type of design.

Chapter 7

Application II: Design and Evaluation of a Mobile Telephone Menu Prototype

7.1 Introduction

The design and evaluation reported in this chapter are motivated by two factors. The first factor is that the design framework developed in chapter 5 has only been followed in a very constrained context. The design presented here is carried out in an optimal context *i.e.*, with no restriction on the type and quality of the sounds. This is justified by the fact that the latest generation of mobile devices can play CD-quality sounds. The second factor was the absence of studies investigating the effect of non-speech sounds on the performance of actual navigational tasks in a large menu hierarchy. The experiment presented in this chapter shows the impact of sounds designed according to the framework described in Chapter 5 on such tasks.

As the technology available while this research was conducted did not allow the designed sonification to be implemented in a real mobile device, the sonification was implemented in a PC-based simulation of a mobile phone menu. Even though computer-based simulations are not as ideal as real prototypes, they have the advantage to enable accurate usability data to be collected.

After the initial description of the prototype simulation, the first part of this chapter will present the sound design process. The main part of the chapter will then be dedicated to the evaluation of this design. The aim of the evaluation is to point out differences between the performance of a set of tasks in "sonified" conditions and in "normal" conditions. This will show whether the sonification increases the usability of the menu or not.

7.2 Test Platform and Simulation

The mobile phone menu used in this study is the NOKIA *6110* menu and a simulation of this device was developed to run this experiment. The simulation was developed in Java version 1.3 Beta; the Java Sound API released in version 1.3 Beta of the Java Development Kit was used to play the audio files that constitute the sonification[1]. Figure 7.1 displays the simulation window for this experiment. For the purpose of this experiment, all the functions of the device were not implemented in the simulation. For instance it was not possible to make a call with the simulation or to access the device phone book. However, it was possible to carry out a number of tasks which made the experiment realistic. These features implemented were:

- Personalising a profile.
- Changing some of the device settings.
- Diverting calls.
- Barring calls.
- Checking and erasing calls.

The best way to become familiar with the platform used in this experiment is to try it out. The simulation that contains all the features required to run the experiment is available on the CD provided with this thesis. Details concerning the use of the simulation can be found in Appendix D.3.

7.2.1 Navigation in the Menu Structure

Only the five buttons required for navigation in the menu were implemented in the simulation. These buttons are represented on Figure 7.1. Button 1 is used to make menu item selections

[1] See the Java web site for more information about Java and the JavaSound API: http://java.sun.com

i.e., go down in the menu hierarchy. Button 2 is used to go back up one level in the menu. Buttons 3 and 4 are used to navigate through a list of menu items at a given level of the hierarchy *i.e.*, move horizontally in the menu[2]. Button 5 allows users to jump to the top of the menu from any node of the hierarchy.

The NOKIA *6110* allows users to jump to a menu item directly by entering the numerical position of the requested node. For example, the menu displayed in Figure 7.1 could be accessed by entering the numbers: 3, 3, 2 and 3 sequentially from the keypad. This feature, clearly targeting expert users, was not implemented in the simulation. Indeed, all the subjects of the experiment were novice users and were not informed of the existence of this function.

7.2.2 Menu Structure

The top of the NOKIA *6110* menu hierarchy is an introduction screen that provides the user with various pieces of information. As far as navigation is concerned, the user is given the option of either going to their address book (*Names*) or selecting the main menu *Menu*. As mentioned earlier, the address book was not implemented in the simulation, therefore the experiment's respondents only had the option to choose *Menu* at that stage, that is, they could only click on the left button (Button 1). Figure 7.2 represents this top-level screen.

The NOKIA *6110* menu contains nine main menus. The structure of these menus is very similar to that of the NOKIA *8210* (see Chapter 6). As Figures 7.3, 7.6, 7.7, 7.8 show, the main menu branches have very different structures from one another. Figure 7.4 and 7.5 also show that the structure of the *Call Register* menu is dynamic. The size and structure of this menu depend on the calls registered in the menu. The main menu items are listed below:

1. *Messages*
2. *Call register*
3. *Profiles*
4. *Settings*
5. *Call divert*
6. *Games*

[2]The interface allows users to loop over a list of menus.

Figure 7.1: Simulation of the NOKIA *6110* menu used for this experiment. The five numbered buttons are clickable. Buttons 1 and 2 allow the operation labelled above them to be carried out (Select and Back on this picture). Buttons 3 and 4 allow navigation through a list of menu options and Button 5 allows users to jump back to the top of the menu from anywhere in the menu.

Figure 7.2: Top level menu item. The time is displayed as well as the name of the active profile: (*meeting*)

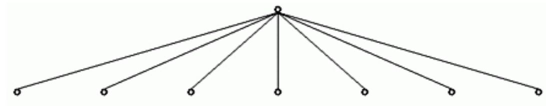


Figure 7.3: Structure of the *Messages* menu.

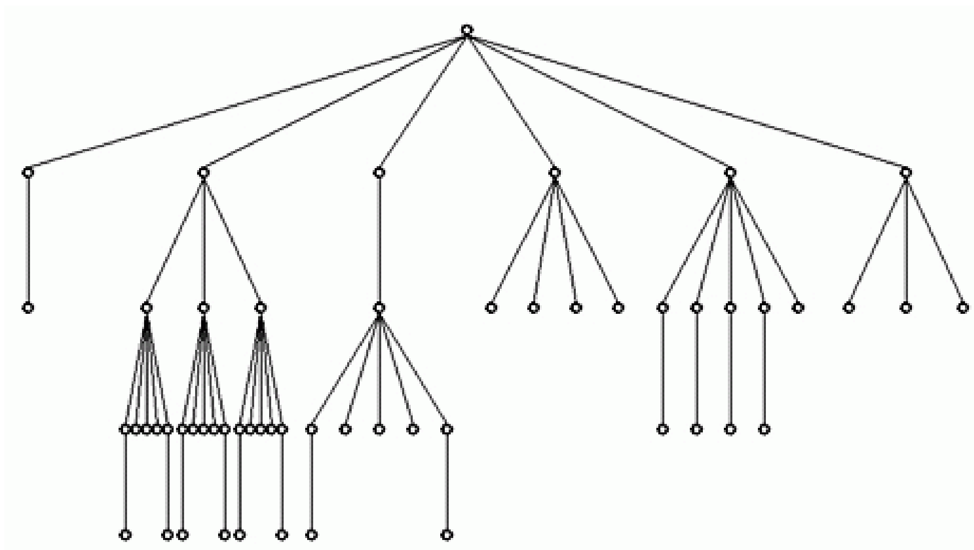


Figure 7.4: Structure of the *Call Register* menu when the menu contains call records.

7. *Calculator*
8. *Calendar*
9. *Infrared*

7.3 Design of the Sounds

7.3.1 Sonification of the Main Menu

According to Principle 1, non speech sounds should be used to increase the semantic differences between the labels of the main menu options. The implementation recommendations of this principle suggest that timbral differences are a particularly efficient way to achieve that goal. Therefore, the most significant menu branches have been sonified with different instruments.

The instruments used for this sonification were all part of the standard sound banks of the En-

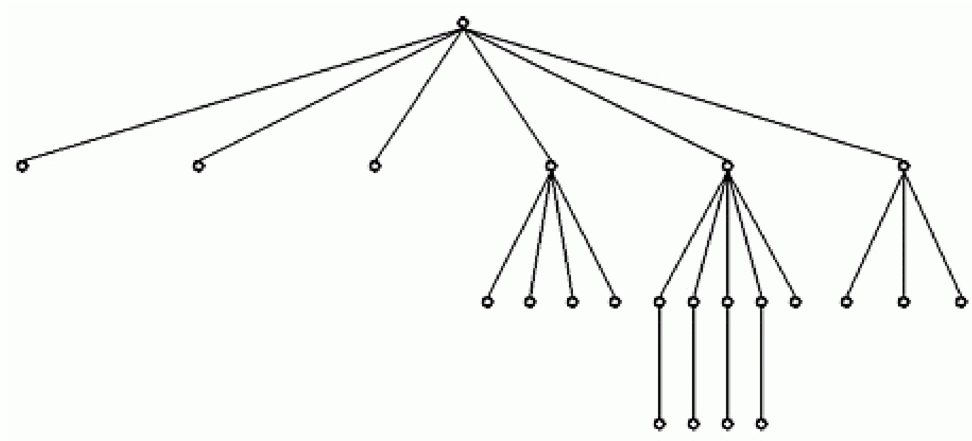


Figure 7.5: Structure of the *Call Register* with no call recorded.

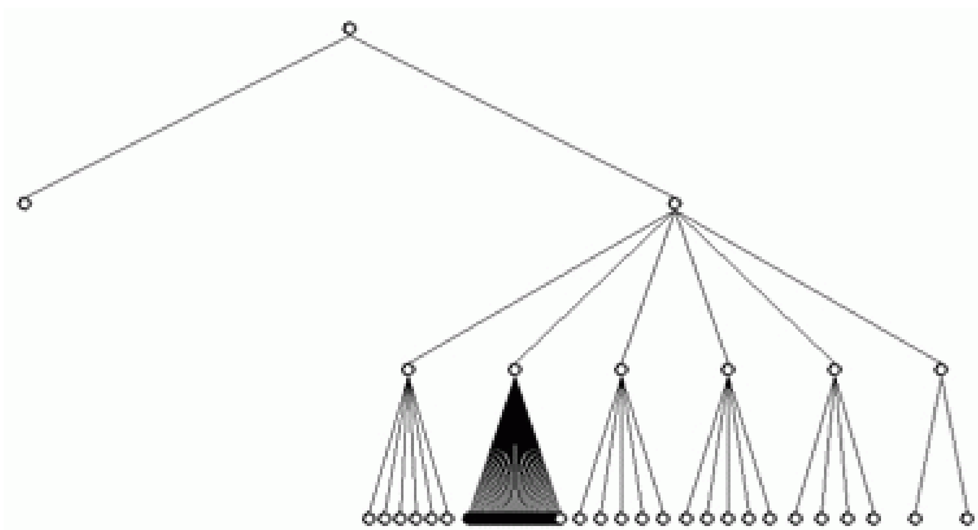


Figure 7.6: Structure of the menu options available for one profile. The *Profiles* menu contains five distinct profiles and therefore five such branches.

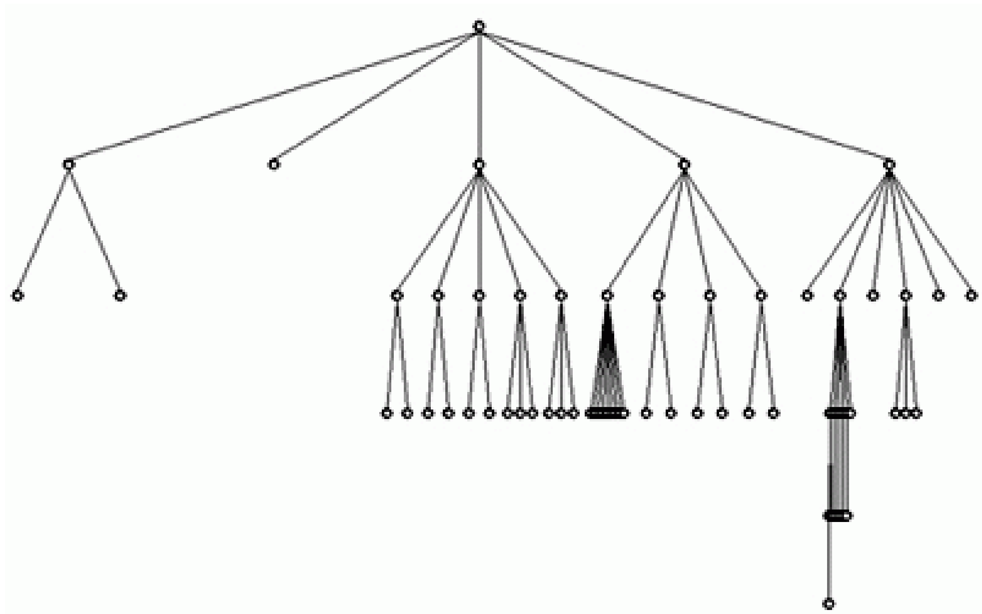


Figure 7.7: Structure of the *Settings* menu.

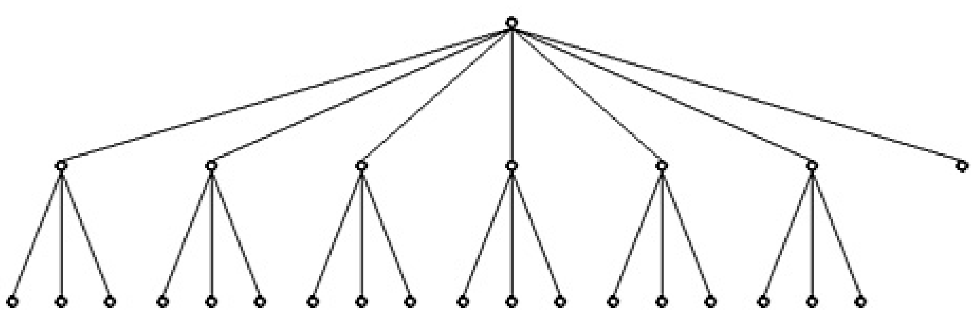


Figure 7.8: Structure of the *Call Divert* menu.

soniq Mr Rack synthesiser. The motifs were first assembled using a MIDI sequencer, Cakewalk Pro Audio 7, and recorded as audio file with Steinberg WaveLab. The recording was done at CD quality, on 16 bits and with a sampling rate of 44.1kHz.

Some of the instruments described in this paragraph are electronic sounds that do not relate to the traditional categories of acoustic instruments. In fact, some of these instruments only exist in the Ensoniq Mr Rack synthesiser. Therefore, the reader is encouraged to listen to the sounds available with the mobile phone simulation. For the first main menu, *Messages* (option and branch), the piano was used. The second main menu, *Call Register*, was sonified with a synthesised combination of a percussive sound, similar to a bell with a *layer*-type timbre. This instrument will be referred to as: *evolution* (name used for it in The Ensoniq Mr Rack synthesiser). The third main menu, *Profiles*, was sonified with a layer-type instrument: *Airy voices* (according to the Mr Rack label). The fourth main menu, *Settings* was sonified with a violin playing pizzicato. For the fifth and last non-trivial main menu, *Call Divert*, a synthesised pitched percussion referred to as *textures* in Mr Rack. Table 7.1 summarises the properties of the instruments used in the design.

As Table 7.1 suggests, the instruments used in this design share some properties as far as envelope is concerned. According to the principles in Chapter 5, the sounds used should be brief or, if they are not, their temporal envelopes should hold certain properties. In particular, in order to make the transition between sounds smoother while navigating, the sustain part of their envelope should always be short, whereas the release or decay parts can last longer.

In addition to the timbral differences, Principle 1 has been applied by using syntactic differences between motifs. Table 7.2 shows that the main menu motifs involve an alternation of chords an arpeggios. This provides the listener with additional cues to categorise the main menus. For instance, the motifs used for *Call Register* and *Call Divert*, designed with timbres that are not very different, are respectively an arpeggio and a chord.

As Table 7.2 also demonstrates, the register of the main menu option motifs have also been chosen to distinguish the motifs further. Again, the register, like the temporal envelope and most of the other parameters of the design have to be manipulated with care, as they have a great influence on the aesthetical quality of the sonification.

Instrument Properties	*Messages*	*Call Register*	*Profiles*	*Settings*	*Call Divert*	*Others*
Name	piano	evolution	airy voices	violin, pizzicato	textures	percu-pitched
Instrument type	percussion	synthesised hybrid	synthesised	plucked string	synthesised percussion	percussion
Attack	sharp	sharp	slow	sharp	sharp	sharp
Sustain	short	short	long	short	short	short
Release	long	long	medium	short	long	short

Table 7.1: Properties of the instruments used for each main menu branch.

Motif Properties	*Messages*	*Call Register*	*Profiles*	*Settings*	*Call Divert*	*Others*
Type *Register*	chord low	arpeggio high	chord medium	arpeggio low	chord medium	arpeggio medium

Table 7.2: Properties of the motifs used to sonify the main menus.

As for the main menu option design of the last chapter, the top level motifs are based on a harmonic progression. In Sample 27 this progression is played with a piano. It involves six different chords, the first five chords are used for the first main menu options while the sixth chord is used as the basis for the design of the remaining four main menus.

Sample 27

Harmonic progression used for the sonification of the main menu.

Sample 28

Reverse of the harmonic progression used for the sonification of the main menu.

One of the requirements of the design involves ensuring that the streams, played while navigating, sound homogeneous whatever path is taken. This implies that the harmonic progression proposed above has to sound right when it is reversed. Sample 28 suggests that this was not achieved in our initial design. As a consequence, the first chord of the sequence was modified in order to make the sequence sound right both ways. The modified sequence is presented in Sample 29. The reverse modified sequence is played in Sample 30.

Sample 29

Modified harmonic progression used for the sonification of the main menu.

Sample 30

Reverse modified harmonic progression used for the sonification of the main menu.

Sample 31 illustrates the final design achieved for the main menu options. The first five sounds of this sequence were used as the basis for the design of the five non-trivial branches of the menu. This design is described in the next sections.

Sample 31

Sequence of the nine motifs used to sonify the main menu options.

7.3.2 Sonification of the Main Menu Branches

The design of the main menu branches will be addressed for each branch in the next sections. This section presents the common design principles used across all the branches and relates them to the Principles proposed in Chapter 5.

Principle 1 As seen earlier, the first principle has been applied by designing each menu option sound with a different instrument.

Principle 2 The feedback to changing level in the hierarchy is consistent in all the branches and is achieve with:

1. The duration of the motifs decreases as the depth of the hierarchy increases.
2. The number of notes of the motifs decreases as the depth of the hierarchy increases.
3. A brief percussive sound which is specific to each level of the hierarchy is played simultaneously to each motif.

Principle 3 The relative position of menu option within a list of option is represented by the position of the motif within the melodic or harmonic progression allocated to the menu.

Principle 4 The number of notes used in a motif allocated to the root of a branch is directly dependent on the size of that branch.

Figure 7.9 illustrated the inheritance mechanism involved in the application of the above rules. The next sections provide additional information on the design of the sounds for each main branch. However the reader is encouraged to listen to the sonification available on the CD provided with this thesis (directions are provided in Appendix D.3).

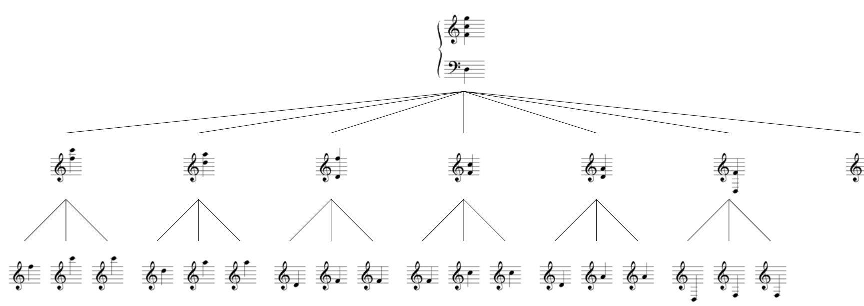


Figure 7.9: Representation of the *Call Divert* menu sonification.

7.3.3 Messages Menu

The implementation of the *Messages* menu in the simulation only contains a limited number of nodes. Therefore, by application of Principle 4, the options of the *Messages* menu have been sonified with very simple motifs. In fact the motifs all contain a single note (See mobile telephone simulation D.3). The main aspects of the design are:

- Motifs were designed with piano sounds

- The duration of the motifs decreases as the depth in the hierarchy increases.
- The number of notes per motif decreases as the depth in the hierarchy increases.
- When played in sequence, the sounds of the options of this menu involve a melody contour which is specific to this menu.

7.3.4 Call Register Menu

The particularity of the *Call Register* menu is that it is dynamic (see Figures 7.4 and 7.5). Its main function being to keep track of calls made and received, the number of nodes in this menu is variable. The sonification has been designed to support this property.

As seen earlier in this chapter, the root of this menu sonification is an arpeggio. The rest of the branch is derived from the root in a similar fashion to that used in Chapter 6, in order to follow Principles 2, 3 and 4. The main aspects of the design are summarised below (Again, listening to the sonification is recommended, see Appendix D.3):

- The instrument used for this menu is: "evolution".
- The motifs used in this menu are all arpeggios or single notes
- The first three options of the menu are descending arpeggios, and the sequence of these arpeggios is itself descending. The last three options of the menu are ascending arpeggios, and the sequence of these arpeggios is itself ascending. This division in two groups reflects the semantic division of the menu option labels. The first three options provide access to phone numbers: *Missed calls*, *Received calls* and *Dialled numbers*, whereas the other three provide a different type of functionality: *Erase lists*, *Call duration* and *Costs*.
- The duration of the motifs decreases as the depth in the hierarchy increases.
- The number of notes per motif decreases as the depth in the hierarchy increases.

7.3.5 Profiles Menu

The profiles menu contains five identical branches: one for each profile available on the device. The motifs used in each branch are all derived from the root of that branch (each top option of the *Profiles menu*). The inheritance rules are the same across the five branches (See simulation in Appendix D.3). The main aspects of the design are:

- The instrument used for the motifs of this menu is: “airy voices”.
- All the sounds of the menu are chords.
- The duration of the motifs decreases as the depth in the hierarchy increases.
- The number of notes per motif decreases as the depth in the hierarchy increases.
- The sequence of chords allocated to the top options of the menu form an ascending motion

7.3.6 Settings Menu

This menu has the most irregular structure of all the main menus both structurally and semantically. As a consequence, keeping the sonification homogeneous was the most challenging aspect of the design (Figure 7.7 illustrates the irregular structure of the menu). Here are the main aspects of the design:

- The instrument used for that branch is a violin playing pizzicato.
- The number of notes in each of the top main options of this menu is equal to the number of submenus these options have (literal application of Principle 4).
- All the motifs of this menu are arpeggios or single notes.
- The contours of the arpeggios used in this menu are all based on the same scale.
- The duration of the motifs decreases as the depth in the hierarchy increases.
- The number of notes of the motifs decreases as the depth increases.

7.3.7 Call Divert Menu

Unlike the previous menu, the *Call divert* menu has a very regular structure (See Figure 7.8). Semantically, all the main options of this menu but the last are very closely related:

1. Divert voice calls without ringing
2. Divert when busy
3. Divert when not answered
4. Divert when phone off

5. Divert all fax calls

6. Divert all data calls

7. Cancel all diverts

Consequently, by application of Principle 1, the last option motif was designed slightly differently from the other option motifs. The first five motifs are two-note chords that follow a descending pattern based on the tonality of the menu. The last one does not follow the harmonic progression of the first 6 motifs and is consequently isolated from the other options. This distinction is however very subtle in order to retain the homogeneity of the sonification. The main aspects of the design are summarised below:

- The instrument used for this menu is: "Textures".
- The motifs used in the menu are all chords or single note motifs.
- The main options of this menu are divided into two semantic categories: the first six option on the one hand and the seventh option on the other hand. The motifs allocated to these menus reflect this division.
- The duration of the motifs decreases as the depth in the hierarchy increases.
- The number of notes per motif increases as the depth in the hierarchy increases.

7.4 Experiment

An experiment was conducted to evaluate the effectiveness of the sonification. The goal of this experiment was to measure the influence of the sonification on the performance of a set of typical navigational tasks. To achieve this, a set of tasks was designed to tackle various aspects of navigation, such as searching and retrieving menu items. Two groups of subjects were recruited to carry out these tasks in a laboratory (see below). The first group carried out the tasks on a simulator in which the sonification designed in the present study was implemented. This group will be referred to as Group 1 in the remainder of this chapter. The second group were asked to carry out the same tasks in a non-sonified simulation. However, the simulation was not completely silent; the sound made by a key-press of the real NOKIA *6110* was recorded and was used as a keystroke cue for this group. The purpose of this was to make the simulation for Group 2 as similar to the original phone as possible. This group will be referred to as Group 2 in the remainder of this thesis. The sonified simulation is available on the CD attached to this thesis[3].

7.4.1 Participants

Two groups of twelve participants from the University of Glasgow, were used for this experiment. The participants, all from the department of Computing Science, were a mixture of undergraduates (two in Group 1 and one in Group 2), postgraduates (eight in each group) and research assistants (two in each group). The twelfth subject of group 1 was an undergraduate from the department of Physical Sciences.

Experience with mobile phones was the main factor taken into account in this experiment. As the navigation review demonstrated (see Section 2.4), it is likely that non-speech sounds affect navigation in a hierarchical structure in both learning and retrieving phases. Consequently, it was important for subjects to have no knowledge of the hierarchy being investigated. Therefore, no experience of using the NOKIA *6110* or any other Nokia mobile phone was a criterion of selection. In addition, to limit the influence of transfer from phones of other brands, subjects with no mobile phone experience at all were given priority in the recruitment of subjects. Ten such respondents were found, while the remaining two per group had limited experience of using a mobile phone menu. As all subjects were computer literate, they were all familiar with navigation in hierarchical structures.

7.4.2 Experimental design

The experiment was entirely implemented as a Java application. Figure 7.10 shows the two windows involved in the experiment. The left window is the phone simulation, in which the participants had to carry out the tasks and the right window contains the instructions that they had to follow. Once a task had been completed, all the participants had to do was press the space bar to display the next instruction.

The simulation only allowed users to complete certain tasks. As mentioned earlier, the five top buttons of the simulated device could be clicked on. The four middle buttons allowed users to navigate in the menu by: going through a list of items using the up and down arrows, selecting

[3]For directions on how to use the simulation, see Appendix D.3

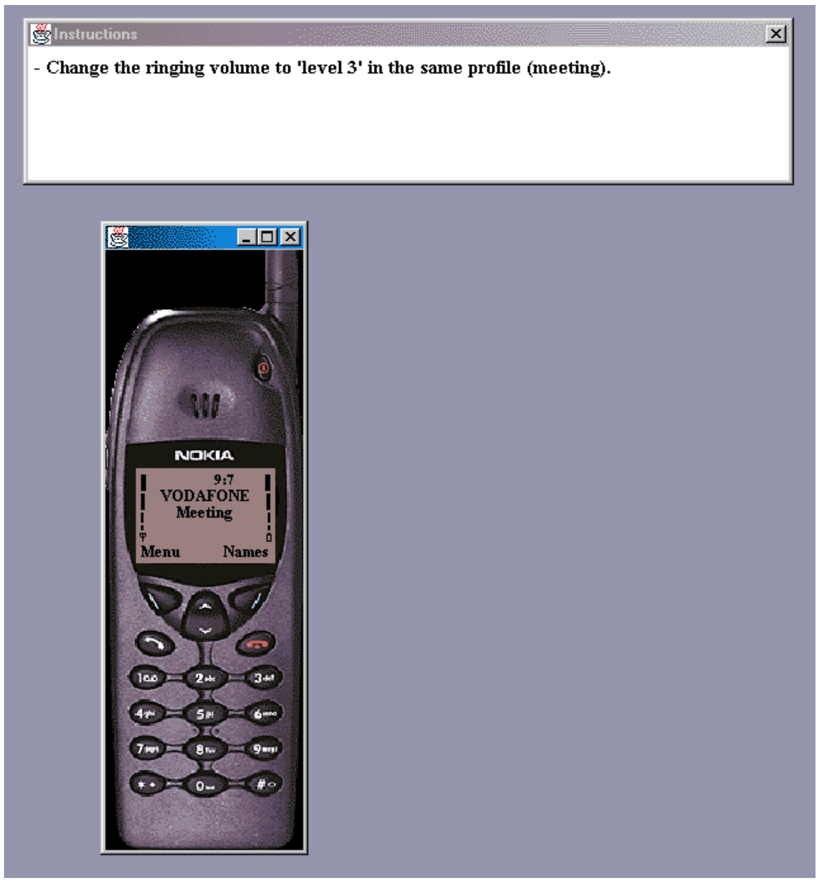


Figure 7.10: Screen-shot of the experiment environment. It involves the simulation window (at the bottom) and the instructions window (at the top).

an item by clicking on the left button or going back by clicking on the right button. In addition to this group of buttons, clicking on the red button at the top-right of the keypad allowed users to jump to the top of the menu from any position in the menu hierarchy. This conforms to the functions of keys on the original handset. The remaining keys were not implemented and, consequently, alphanumeric input was not made possible. Alphanumeric input however was not necessary as our interest was in menu navigation, not in text input. A large number of operations are available using only five implemented keys. Therefore the simulation not only allowed users to navigate within the menu, but also to carry out operations such as changing the parameters of the device, checking and erasing a call made or received. This functionality allowed the formulation of many typical tasks in terms of *operating* functions of the phone.

In order to assess the effectiveness of the sounds used to support navigational tasks, a list of fifty six tasks has been designed. The choice of tasks was based on the review of navigation carried out in Chapter 2. These tasks allowed us to tackle three different categories of search tasks and additional retrieving scenarios:

1. Search tasks in which the path to take is explicit in the task formulation. For example, "Activate the outdoor profile" is a task whose formulation contains the information needed to complete it. The mental matching process represented in Figure 2.6 of Chapter 2 illustrates this.

2. Search tasks in which the path to take is not explicitly formulated in the task, for example, "Activate the call barring service for all incoming calls". The mental matching process illustrated by Figure 2.7 of Chapter 2 demonstrates this.

3. Search tasks involving *isolated* menu items. Isolated menu items are items with a label that does not relate semantically to the branch in which it is located. *Alarm Clock* for example, is a label that does not relate to the branch to which is belongs to (*Phone Settings*), as opposed to *Divert All Calls*, which is strongly semantically linked to its parent (*Call Divert*). Therefore, items such as *Alarm Clock* are referred to as isolated menu items.

4. Retrieving explicitly formulated items.

5. Retrieving non-explicitly formulated items.

6. Retrieving isolated menu items.

The first seven tasks were considered as practice tasks to ensure that all the subjects understood how to navigate in the menu in the Java simulation of the NOKIA *6110*. Among the 49 other tasks, there were 6 occurrences of the following task: *Press the red key to go back to the top level menu*. This task was a means to ensure that all participants would start the subsequent task at the same location in the menu and to make them familiar with this functionality. All participants were free to use this key to jump to the top of the hierarchy at any point of the experiment.

One task (Task 49) differed from the other tasks for it pertained to a menu item that did not exist in the menu: *Is there a "Template" submenu in the "Messages" menu?*. Consequently, this task could not be included in the analysis of the remaining 42 tasks.

The other 42 tasks, referred to as *trial tasks* in the rest of this chapter, involved menu items spread in menu branches according to the figures presented in Tables 7.3, 7.4 and 7.5.

These tables involve the three different types of tasks described above: search tasks with explicit paths, search tasks with non-explicit paths and search tasks involving isolated menu items. Table 7.3 contains explicit tasks involving the *Call Divert*, *Profiles* and *Call Register* menus. The tasks listed in Table 7.4 are non-explicit tasks, they all involve items from the *Call Barring* menu, contained in the *Settings* menu. Finally, Table 7.5 involves tasks related to isolated menu items, all items were located in the *Settings* menu. The complete list of tasks is available in Appendix C.

Additional tasks for Group 1

In order to gather additional information concerning how the sounds were understood by the subjects, Group 1 were asked to carry out two more tasks than Group 2 in blind conditions *i.e.*, with no visual feedback at all. After they had completed the 56 main tasks and filled in the workload form, the mobile phone screen went blank and they were asked to try and complete the following tasks relying only on the audio feedback:

1. Press the red button to go back to the top level. Then try and activate the last profile in the "Profiles" menu.

2. Try and cancel all call diverts (you may use the red button to go back to the menu first). If you do not know how to do that, just try and select the "Call divert" menu.

Call Divert	*Profiles*	*Call Register*
18	8	23
20	9	34
32	11	35
37	12	36
38	14	42
44	16	57
56	24	
	33	
	40	
	50	

Table 7.3: Occurrences of tasks related to the three main menus: *Call Diverts*, *Profiles* and *Call Register*. Each column contains the number of a task that relates to that column's menu. For instance, 7 tasks involve the *Call Divert* menu, and these tasks are Tasks number: 18, 20, 32, 37, 38, 44, 56.

Call Barrings
28
30
43
51
55

Table 7.4: Occurrences of tasks related to the *Call barrings* sub-menu.

Welcome Note	*Time*	*Language*	*Any Key Answers*	*Own Number Sending*	*Speed Dialling*
21	22	27	25	26	31
52	39	41			53
	48				

Table 7.5: Occurrences of tasks related to isolated menu items in the *Settings* menu

These tasks were designed to collect additional information on how subjects associate sounds with main menu items. To complete the first task, a subject is required to know which sound is associated with the *Profiles* menu. In addition, the subjects needed to locate the last option of the *Profiles* menu.

The second task had similar objectives to the first. However, to complete this task subjects were required to know that 'Cancel all' was the last sub-menu in the 'Call divert' menu.

Workload

At the end of the experiment, each participant was asked to rate the following items of a NASA TLX workload form on a graphical scale: Mental Demand, Physical Demand, Time Pressure, Effort Expended, Performance Level Achieved, Frustration Experienced. In addition, the item "Annoyance Experienced" was added to the questionnaire. This item is often added to the workload questionnaires used by Brewster and colleagues [Bre97] to answer the criticism that sounds are annoying.

7.4.3 Hypotheses

The principal hypothesis of this experiment was that navigational tasks would be performed more efficiently with auditory feedback. In order to assess this qualitative hypothesis, a number of quantitative hypotheses have been postulated.

The main quantitative hypotheses of this experiment were concerned with the number of key-presses and time taken to perform the tasks. Non-speech sound is hypothesised to help users build a better mental model of the menu structure. Consequently, a better global mental representation of the hierarchy through sounds should help Group 1 complete the tasks more effectively overall. On the other hand, assuming that non-speech sounds help encode and recall the menu items, the performance of Group 1 should improve more noticeably throughout the experiment than the performance of Group 2. The hypotheses of the experiment are listed below:

1. **Number of key-presses to complete the tasks**

    (a) **Overall difference between groups**
    Globally, Group 1 is hypothesised to complete the tasks with less key-presses than Group 2.

(b) **Evolution of the difference between groups**

In addition, it is hypothesised that the improvement of performance during the experiment will be greater for Group 1 than for Group 2.

2. **Key-press time**

   In a previous study, where non-speech sounds were used along with speech, a significant increase of the time taken to complete the tasks was noticed [Bre97]. This can be justified by the fact that non-speech sounds were played before speech sounds, which increased the duration of the interaction. On the contrary, in this experiment sounds are played simultaneously to users' actions. Given the care taken to minimise the length of the sounds, it is expected for the average key-press time not to be significantly different for both groups.

3. **Errors**

   As some of the tasks are demanding, it is likely that some subjects will fail to complete them. Failing to complete a task involves either completing a task incorrectly, or giving up on the task. It is also hypothesised that Group 1 will complete more tasks successfully than Group 2.

7.5 Results

For each task and subject, all the menu selections were recorded, as well as the duration of time between two consecutive menu selections. This information allows a complete picture of the subjects' performances to be drawn, and for the hypotheses of this thesis to be confirmed or denied.

7.5.1 Number of Key-Presses

Firstly, the total number of key-presses for each task and participant over the whole experiment was accounted for. This analysis provided a first indication of the differences of performances between the groups. Subjects from Group 1 averaged 1162 key-presses to complete the experiment, whereas subjects from Group 2 took 17% more, that is, 1362 key-presses. There was a difference between the variances of the results within the groups: 34074 for Group 1 and 108038 for Group 2. An F-test on both samples showed that indeed, the total number of key-presses of Group 2 subjects was significantly more dispersed than that of Group 1

($F_{23} = 0.31, p = 0.034$). Overall the variance was high, which was to be expected due to the limited number of subjects and the nature of the tasks involved. Consequently, a t-test showed that the difference of performance in terms of number of keystrokes did not achieve significance ($t_{11} = -1.83, p = 0.084$).

The number of key-presses is a good indicator of how efficiently the tasks have been performed. However, several issues arise whilst running statistical tests on this value:

- The number of key-presses taken to complete a task does not take into account whether a task has been completed successfully or not.
- The distribution of the total number of key-presses is heavily tailed, especially for Group 2.
- The numbers of key-presses vary significantly according to the task being completed.

As a consequence the raw data collected during the experiment did not meet the requirements of classical statistical tests in order to provide meaningful information. It was thus necessary to devise alternative sets of data based on the raw data to deal with the ill-distribution of the raw data and to account for the completion of tasks in this analysis.

In order to deal with the tailed distribution of the data (Group 2 especially), a common practice in statistical analysis involves applying adequate mathematical functions, such as *log* and *squareroot*, to the data [How92]. However, *log* or *squareroot* transformations would not be appropriate in this instance since the data of Group 2 are heavily tailed in both low and high ends. An alternative would involve using trimmed samples, but given the relatively small amount of data, reducing it further would not be appropriate either. In addition to this, all these transformations do not address the issue of whether the tasks were completed successfully or not. In the present case, the most adequate transformation of the data is a normalisation of the number of key-presses for each task and subject. This method allows success or failure in completing a task to be taken into account: For each task, a value between 0 and 1 is allocated to each subject, this value is calculated using the number of key-presses for the task and subject divided by the maximum number of key-presses over the 24 subjects for this task. This process is illustrated by the data represented in Tables 7.6 and 7.6. Table 7.6 shows the initial data for a particular task (Task 52), and Table 7.7 shows the normalised data for the same task.

	Group 1	*Group 2*
subject 1	8	36
subject 2	15	23
subject 3	6	19
subject 4	22	15
subject 5	12	12
subject 6	12	33
subject 7	21	9
subject 8	14	13
subject 9	9	6
subject 10	18	0
subject 11	14	10
subject 12	9	18
total	160	194

Table 7.6: Number of key-presses taken by the subjects of groups 1 and 2 to complete task 52

The distribution of the normalised number of key-presses for each group, referred to as NNKP in the remainder of this chapter, is far less skewed than the raw data obtained and the variances of the two groups are in the same range. This refined set of data enables a more meaningful statistical analysis to be performed. In addition, this new set of data will allow the task completion success to be easily taken into account later in this chapter.

The variance of the total NNKP per group across the 42 trial task is 2.46 for Group 1 and 2.31 for Group 2. A t-test on these 2 groups of 42 values shows that the difference of means (4.07 for Group 1 versus 4.70 for group 2) approached significance ($t_{41} = -1.84, p = 0.069$). Similarly, the variances of the total NNKP per task for each group were quite similar (3.43 for group 1 versus 8.40 for group 2). A t-test on these 2 groups of 12 values shows a significant difference between the means (14.25 for Group 1 versus 16.42 for Group 2, $t_{11} = -2.19, p = 0.040$).

Regardless of whether the tasks have been completed successfully or not, these results confirm the hypothesis concerning the total amount of key-presses. Most of the tasks have been completed successfully, however, the unsuccessful completion of tasks is an important way to

	Group 1	*Group 2*
subject 1	0.2	1
subject 2	0.4	0.6
subject 3	0.2	0.5
subject 4	0.6	0.4
subject 5	0.3	0.3
subject 6	0.3	0.9
subject 7	0.6	0.3
subject 8	0.4	0.4
subject 9	0.3	0.2
subject 10	0.5	0
subject 11	0.4	0.3
subject 12	0.3	0.5
total	4.4	5.4

Table 7.7: Normalised values from Table 7.6. The value of easch cell of this table is equal to the value of the corresponding cell in Table 7.6 divided by the maximum value (36) of Table 7.6

measure the success of the use of auditory input in navigation of a menu hierarchy. In Section 7.5.3, the number of errors will be analysed in detail, but it is also important to take them into account as part of the present analysis.

There are two reasons why it is important to take errors into account in this analysis: a task might be failed because a participant may have given up looking for the right menu item after a long time spent searching. In this case, where the participant is lost, the number of keystrokes for this participant and task is likely to be high. Also, the re-occurrence of a task that has not been completed successfully by a subject at a previous stage may result in the subject not attempting it again. In this case, the total number of key-presses for this task and consequently the group to which the participant belongs, would be low. Depending on the context, unsuccessful tasks may then result in an unhelpful deviation of the data. The operations described here must be taken as what they are: an attempt to understand more about the participants' performance.

There is no ideal way to solve this problem, but it is arguable that the following treatment is a fair adjustment of the data: this treatment simply involves replacing the number of key-presses for a given failed task, by the maximum number of key-presses taken by any of the 24 participants to complete that task successfully. For example, in Table 7.6 and in Table 7.7, participant 10 from group 2 scored 0. This means that he/she did not attempt to complete the task at all, because he/she had failed it previously. Using the treatment described above, this value will be replaced by 36 in Table 7.6 and 1 in Table 7.7. As an example, Table 7.8 and Table 7.9 show how this treatment applies respectively to Table 7.6 and Table 7.7.

The same statistical tests as above have been performed on these data. Again, the variance of the total NNKP across the 42 analysed tasks were similar for both groups (3.48 for group 1 and 4.04 for group 2). A t-test showed that there was a significant difference between the means of total NNKP per task of the 2 groups ($t_{41} = -2.26, p = 0.027$). Similarly, a t-test showed that there was a significant difference between the means of total NNKP per subject of the 2 groups ($t_{11} = -2.24, p = 0.040$).

These results show that taking into account success or failure to complete a task increases the differences between group 1 and group 2 as far as numbers of key-presses are concerned. In addition, with adjustment of the data, Hypothesis 1b is validated.

	Group 1	*Group 2*
subject 1	8	36
subject 2	15	23
subject 3	6	19
subject 4	22	15
subject 5	12	12
subject 6	12	33
subject 7	21	9
subject 8	14	13
subject 9	9	6
subject 10	18	*36*
subject 11	14	10
subject 12	9	18
total	160	230

Table 7.8: Data adjustment in order to take task completion into account in the analysis of keystrokes number: Number of key-presses taken by the subjects of groups 1 and 2 to complete task 52

	Group 1	*Group 2*
subject 1	0.2	1
subject 2	0.4	0.6
subject 3	0.2	0.5
subject 4	0.6	0.4
subject 5	0.3	0.3
subject 6	0.3	0.9
subject 7	0.6	0.3
subject 8	0.4	0.4
subject 9	0.3	0.2
subject 10	0.5	*1*
subject 11	0.4	0.3
subject 12	0.3	0.5
total	4.4	6.4

Table 7.9: Normalisation of the data from Table 7.8

Evolution of the Performance of the Groups in Terms of Key-Presses

The second hypothesis that was formulated regarding key-presses is related to the evolution of the respondents' performance over time, that is, for the duration of the experiment. As for the previous analysis, the original data, and the treated data were analysed. The method used to compare how the difference between groups evolved, involved running a linear regression on the following variable for the 42 trial tasks:

$$t(task) = t_1(task)/t_2(task) \tag{7.1}$$

where $t_1(task)$ and $t_2(task)$ are the total number of key-presses for $task$, for Groups 1 and 2 respectively. Similarly, $t_1^a(task)$ and $t_2^a(task)$ will be the notation for the related adjusted variables.

Without adjustment, the slope of the regression line is almost flat, which indicates that the difference of performance of the 2 groups is constant throughout the experiment (see Figure 7.11). However, as has already been mentioned, this measurement is highly biased by the fact that the

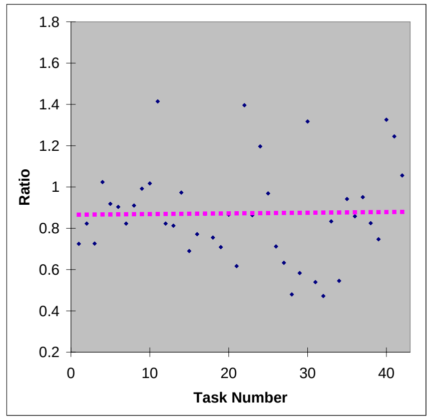


Figure 7.11: Regression line through the set of values of $t(task)$ for the 42 trial tasks

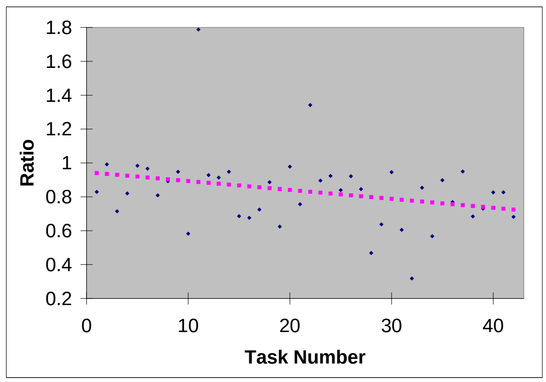


Figure 7.12: Regression line through the set of values of $t^a(task)$ for the 42 trial tasks

number of failed tasks increased more in Group 2 than in Group 1 throughout the experiment (see Section 7.5.3). Most of the late experiment tasks that were failed have not been attempted at all and the number of key-presses for such tasks and subjects is zero. Therefore, the analysis performed on the raw data would be misleading.

With adjustment, the regression on $t^a(task)$ for the 42 tasks showed a big increase in the difference of performance of the 2 groups throughout the experiment (see Figure 7.12). The initial value of the regression line ($task = 1$) is 0.94, versus 0.72 for the last task ($task = 42$). In other words, group 1 took 6% less key-presses than group 2 at the beginning of the experiment, and 28% less at the end. Once more, these figures validate the hypothesis 1b.

7.5.2 Key-Press Time

The main concern with the use of non-speech sound is that it tends to slow down the interaction. In this experiment, however, sounds have been designed to coincide with the time spent by a user to visualise a menu item. The comparison of the average key-press times for each group will indicate whether the design has been successful.

The time spent between two key-presses has been collected for each task and participant. During the whole 56 tasks that constitute the experiment, the mean time spent on a menu item was $1.14s$ for group 1 versus $1.08s$ for Group 2. A t-test on the average key-press time per subject showed that this difference was far from significant ($t_{11} = 0.66, p = 0.51$). A t-test on the average key-press time per task confirms this result ($t_{55} = 74, p = 0.46$). Therefore these results also validate hypothesis 2.

7.5.3 Errors

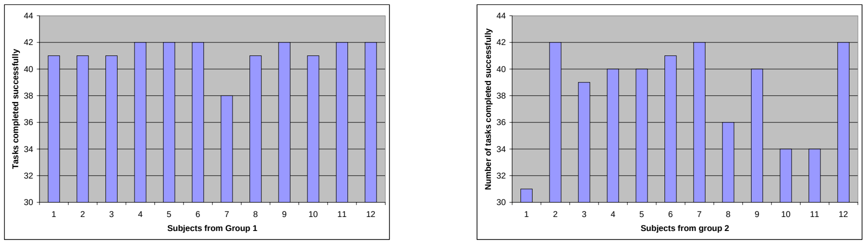
The experimental procedure did not require users to succeed in completing one task in order to proceed to the next task. Therefore, it was to be expected that some participants would give up tasks after a while spent trying to complete them unsuccessfully. This is often the case with real mobile telephone usage Users cannot find the functions they want in the huge number of items available. As seen earlier, there were two different categories of tasks:

1. Tasks where the subjects had to *find* a menu item.

2. Tasks that required the subjects to *complete* an action.

For the first category, a task was considered successful when a subject had selected the required item. For the second category, a task was considered successful when the action was completed.

The average score for Group 1 is 41.25 versus 38.42 for Group 2, respectively 98.2% and 91.5% of the tasks completed successfully on average. A t-test indicated that this difference was significant ($t_{11} = 2.52, p = 0.026$). Figure 7.13 shows that the distributions of Group 1 and Group 2. These statistics validate hypothesis 3.



(a) Group 1 (b) Group 2

Figure 7.13: Number of successful tasks for each group.

Evolution of the Number of Tasks Completed Successfully for Groups 1 and 2

In section 7.5.1, it was mentioned that the number of errors increased more for Group 2 than for Group 1 during the experiment. To prove this point, a linear regression was performed on the differences of tasks completed successfully in each group. The graphical representation of this regression is displayed on Figure 7.14. The correlation between X and Y is close to average ($r = 0.40$). The initial value of the correlation line ($task = 1$) was 0.05, versus -1.67 for the final value ($task = 42$). An ANOVA analysis [How92] showed a high significance of the regression ($p = 0.090$). These results show that the evolution of the difference of performance between the groups in terms of errors made is largely in favour of group 1.

7.5.4 Additional Tasks for Group 1

As mentioned earlier in this chapter, group 1 were requested to perform two tasks in non-visual conditions at the end of the experiment (see description in 7.4.2). For technical reasons, the first subject of the group did not perform these additional tasks. Therefore, only 11 sets of data were collected for these two tasks.

This part of the experiment has been developed for purely exploratory purposes; no explicit hypotheses were formulated concerning these tasks. Therefore, at this stage of the analysis, the type of errors made by the respondents are of interest rather than a detailed statistical analysis of them. Out of the 11 respondents, only two managed to complete the first task accurately whereas five managed to complete the second task correctly.

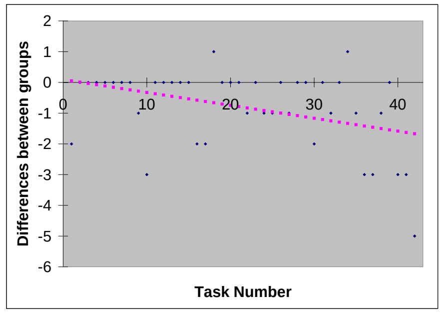


Figure 7.14: Evolution of the differences between group 1 and group 2 in terms of tasks completed successfully

Some interesting points arise if the performance of these tasks is considered. In order to complete Task 1, a respondent had to perform the following actions:

1. Click on the red button to go to the top of the menu. The feedback for this action is the sound associated with the top menu item.

2. Click on the left button to go down one level *i.e.*, to the first of the main menu items. The feedback for this action is a piano chord, which is the sound associated with the *Messages* menu.

3. Use the up or down keys to scan the list of main menu items until *Profiles* has been found. The feedback is the sounds associated with each main menu item whilst being browsed.

4. Click on the left button to select that menu. The feedback is the sound of the first sub-menu item.

5. Use the up and down keys to reach the last sub-menu. There are two main strategies to complete this action:

    (a) Press the up key once. From the first item in the list this action would indeed take the user to the last item in the list.

    (b) Press the down button until a loop has been completed. When browsing the menu item downwards, a series of decreasing pitch sounds are played. This feedback is sufficient to indicate when a loop has been completed.

6. Press the left button to select the last sub-menu, which is *Pager*, if the user has not made any mistakes so far. The feedback is the sound associated with the first sub-menu of the *Pager* sub-menu i.e., *Activate*.

7. Press the left button to select *Activate*. The feedback for this action is the completion sound being played.

Seven out of eleven respondents completed up to step 6 successfully, two of them also completed step 7, but the remaining five stopped after completing step 6. It is likely that they believed the profile would be activated merely by selecting it. Therefore, these 5 subjects did not notice the absence of the completion sound to realise they had not completed the requested action. Among the remaining four subjects, three failed at step 3. Two selected the *Call divert* menu and then performed the next steps successfully, and 1 selected the *Settings* menu. These 3 individuals failed to identify the characteristic sound of the *Profiles* menu. The last subject failed at step 5, and selected the wrong profile.

In a similar fashion, the second task could be deconstructed as follows:

1. Click on the red button to go to the top of the menu. The feedback to this action is the sound of the top menu item.

2. Click on the left button to go down one level *i.e.*, to the first of the main menu items. The feedback for this action is a piano chord, which is the sound of the *Messages* menu.

3. Use the up or down keys to scan the list of main menu items until the current item is *Call divert*. The feedback is the sound associated with each main menu item whilst being browsed.

4. Click on the left button to select that menu. The feedback is the sound of the first sub-menu item.

5. Use the up and down keys to reach the last sub-menu. Managing this presupposes that the user remembers the *Cancel All* menu is the last of these sub-menus. There are two main strategies to complete this action:

   (a) Press the up key once. From the first item in the list, this action would indeed take the user to the last item in the list.

(b) Press the down button until a loop has been completed. When browsing the menu item downwards, a series of decreasing pitch sounds are played. This feedback is sufficient to indicate when a loop has been completed.

6. Press the left button to select the *Cancel all* menu. The feedback for this action is the completion sound being played.

Five out of eleven participants managed to complete the task successfully. Four subjects completed the task up to step 4 then failed to go any further (2 of them) or got lost in the *call divert* menu before giving up (the other 2). This indicates that these subjects did not recall that the *Cancel all* sub-menu was the last of the *Call divert* sub-menus. The remaining 2 subjects of the group did not manage to select the *Call divert* menu.

The interpretation of these results should be treated carefully. What these tasks indicate is not how the sounds have been understood by the respondents, but how well these sounds are understood on their own, in the absence of a visual context. It is important to remember that the sounds have been designed to enhance the semantic content of menu items, not to replace the visual items. Consequently, navigating with sounds as the only feedback can be considered to be a highly demanding task. With this in mind, the success of users ability, even partially, to complete a task in the absence of any visual feedback tends to suggest that the semantic content of auditory cues greatly enhances menu navigation.

Overall, 77.3% managed to select the correct main menu with audio feedback only. A level below, 54.5% managed to select the correct sub-menu: 7 out of 11 when the position of the item was specified (first task), and 5 out of 11 when the position was not specified (second task). However, 5 subjects failed to notice the absence of the completion sound in the first task, which would have indicated that the task was completed.

The results of the workload questionnaire are compiled in Table 7.10, which shows that no significant difference was found between both groups. The most interesting parameter of the workload test was the annoyance of the sonification. The average value of annoyance reached 34.7% (on a scale of 0 to 100). To put this value into perspective, the physical was rated on average at 20%, and the experiment was not physically demanding. In addition to this, annoyance received a lower rating than 'time pressure' did. Given that the users did not have any time constraints to complete the experiment, one can say that the annoyance rating is low.

	Group 1	Group 2	T-Test (p value)
Mental Demand	64.5	51.3	0.185
Physical Demand	19.9	20.5	0.940
Time Pressure	38.2	53.8	0.155
Effort Expanded	64.1	51.3	0.227
Performance Level achieved	54.9	51.3	0.710
Frustration Experienced	46.2	49.5	0.804
Annoyance	34.7	-	-

Table 7.10: Results of the workload test. The last three columns are respectively: the means (in percent) for each component of the workload for Group 1, for Group 2, and the probability ($p(t_{11})$) associated with a t-test that indicates whether the differences of means between the groups are significant or not.

In particular, six out of twelve subjects rated the annoyance of the sounds below 20% whereas one subject only rated it above 80%.

7.5.5 Subjective Ratings of the Sounds

In this type of study, statistical analysis draws a clear picture of the respondents' performance. However, in human computer interaction studies, more informal and subjective comments from the respondents are also very useful in assessing the qualities and defaults of a design. The objective of the questionnaire submitted to Group 1 at the end of the experiment was to gather some feedback on their subjective appreciation of the sounds, from a different angle from that of a workload form. The questions asked to the members of Group 1 were:

- **Do you think that your overall performance benefited from the sounds?**
- **Were the sounds meaningful to you?**
- **What kind of information was provided by the sounds?**
- **What did you dislike most in these sounds?**
- **What did you like in them?**
- **Any other comments?**

The answers of each participant are compiled in Appendix C, but the main points of their answers are summarised below. These are the main reasons for which they perceived the use of sound as a positive aspect of the interface:

- They (the sounds) help remember the different profiles.
- They help notice when one loops over a menu list.
- "you can tell quicker that you are looking at something you've already looked at".
- They tell whether there are submenus under my current position.
- They give feedback to keystrokes (mouse clicks in the experiment).
- They tell if a task was completed successfully.
- They provide a rhythm to selecting menu options.
- the sounds made the experiment more fun than it would have been without the sounds.

Of course there were also negative comments, as the participants were asked what they disliked most about the sounds. The answers included:

- "Some of them (the sounds) started to annoy me."
- "They don't relate to the menu option or I didn't pracitce enough to learn them to the menu options."
- "Some were rather intrusive".
- "Some provided no info (e.g. in change tone). Not immediately apparent why the sounds differ in main menu. No consistency in their use.

7.6 Conclusions

This chapter has presented the complete design and evaluation of a large mobile telephone menu sonification. The motivation of this work was two-fold:

1. Demonstrate that the principles proposed in Chapter 5 could be applied to the sonification of a large mobile telephone menu.

2. Evaluate the effectiveness of the sonification in supporting actual navigation tasks in the menu.

The analysis of the results collected have suggested that the sonification was beneficial to the users in many respects: The users completed the tasks of the experiment more successfully with sound and their performance improved more significantly over the experiment. The main results of the experiment are summarised below:

1. **Number of key-presses to complete the tasks**

   (a) **Overall difference between groups**
   The statistical analysis of the raw data collected did not demonstrate a significant difference between the two groups investigated. However a closer look at the data revealed that the raw number of key-presses was not a meaningful data set because it was biased by whether tasks had been completed successfully or not. The adjustment of the data proposed to overcome this bias showed that, by taking into account success or error, the difference of performance between the two groups was significant (in favour of Group 1).

   (b) **Evolution of the difference between groups**
   From the beginning of the experiment to the end, Group 1 improved their performance more than Group 2.

2. **Key-press time**
   There was no difference between the average time between two keystrokes between the two groups. The sounds did not slow the users down.

3. **Errors**
   Group 1 completed more tasks successfully than Group 2.

In conclusion, this chapter has demonstrated that the design framework proposed in Chapter 5 has allowed for the successful design of a large mobile telephone sonification. The directions provided by this design framework has been beneficial to the design in two respects:

1. The focus of the design was on designing sounds that would improve actual navigation tasks, rather than simply hierarchical sets of sounds. Chapter 3 showed that the previous attempts at designing hierarchical sounds did not base their design on actual tasks

and therefore never demonstrated that such tasks could actually be improved by their sonification [Bre98].

2. The aesthetic character of the sonification has been addressed. The principle of moderation stated in Chapter 5 allowed a homogeneous sonification to be achieved. The subjective feedback collected indicates that most of the participants found the sonification pleasant.

Chapter 8

Practical Sound Design Issues

8.1 Introduction

The present thesis has dealt with a number of theoretical and practical sound design problems. In Chapters 4, 6 and 7 in particular, hierarchies of sounds have been created, raising practical questions regarding the most appropriate means to design these sounds. For small hierarchies such as that dealt with in Chapter 4, there was no real practical design issues, but for the larger hierarchies investigated in Chapters 6 and 7, designing sounds proved a more technically challenging process. In particular, one of the problems that arose was: How can the changes made in a sound be "propagated" to other sounds automatically in order to respect the hierarchical relationships existing between the sounds? This question, as well as others reviewed in this chapter, pleaded in favour of the development of a tool which makes, auditory menus and auditory interface prototyping in general, easier.

The term "sound design" is often used to denominate the creation of *special effects* for games or films. In this chapter, sound design is referred to as the general operation of creating sounds for a given system. In sound design for films, the emphasis is placed on the relationship between the visual and the auditory streams to achieve affective effects. Creating sounds for games involves a similar goal in a context in which the user has an input on the state of the system (there is interaction). In a hierarchical menu sonification, the desired effect is more informative than affective.

In sound design for films, the challenge is essentially compositional. The task of the sound designer / composer is only constrained by predefined time events and intervals. Therefore, a tool like Macromedia Director is perfectly adequate for the job. As far as games are concerned, the same constraints exist: part of the creation of the sonic environment of a game involves the creation of one or more sound tracks. Sound tracks are played in the background of the game. In addition, a game's auditory scene features "special effects" which are sounds generated in real-time or played back to produce feedback to the interaction. From the design point of view, there is a limited number of these sounds or effects and they are not strongly related to one another. Once created and implemented in the game, they can easily be refined independently and re-implemented. On the contrary, the sonification of a large menu, such as those dealt with in the present thesis, involves a much more demanding prototyping process. No appropriate tools are available to carry out this task.

Prototyping tools have proved very successful for the development of interactive systems, because they enable the designer to go through the design and implementation loop, easily and quickly. There is a large variety of general and more specific tools available for graphical design, but no appropriate solution for the design of interactive systems involving audio exists. Designers have the choice between traditional prototyping tools such as Macromedia Director in which the audio functionality is very limited, and, on the other hand sound design tools like Opcode Max[1]. Max has been developed for interactive music, and linking it to a user interface requires a lot of code to be written, in order to achieve a functional prototyping environment. Besides, even though it is a powerful tool, using Max is far from being easy.

The Auditory Display research community has developed a number of tools, but these often only serve a very specific purpose. A large portion of these tools are dedicated to data sonification, or to the creation of virtual sonic environment, which are not relevant to the problem tacked in this thesis. A representative example of the former category is MUSE, a data sonification toolkit bby Lodha and colleagues [LBH+97]. The CAITLIN environment addresses the use of structured non-speech sounds, but focusses on using such sounds in programs [AV97]. Moreover, CAITLIN does not enable interactive sound design and auditory menu prototyping.

The lack of dedicated tools for audio-visual design is justified by the fact that this is a relatively new area of work. However, the design process involved in Chapter 7 has suffered from the lack

[1]Information about MAX and its public domain variant (PD) both developed by Miller Puckette can be found on Miller Puckette's homepage: http://www-crca.ucsd.edu/~msp

of such tools. The menu investigated in that chapter was constituted of several hundreds of nodes and the same amount of sounds had to be designed. Because the sounds had structural relations with each other, and with the hierarchical structure itself, the sounds could not be designed individually. In a design of this scale, it would have been extremely useful that the modifications made on one sound propagate to other related sounds. In order to pursue the research carried out in Chapter 7, the development of a prototyping tool was undertaken. As a result, the tool developed was used for the sonification of the mobile phone prototype reported in Chapter 6.

The present chapter does not only describe *AIDE*, the tool that has been developed, it proposes a global formalism embedding the creation of auditory objects and the integration of these objects in an interaction mechanism. This formalism can be applied to the sonification of menus, but is also valid for more general purposes, as the next sections will demonstrate.

8.2 A Richer Formalism for the Development of Auditory Interfaces

Most commercial prototyping tools (such as Macromedia Director) feature a very basic auditory language. Typically, they enable an audio or a MIDI file to be played or stopped when an interaction event occurs. The objective of this section is to present the formalism that is required in order to undertake the sonification of hierarchical menus. This is best explained with the following example: Figure 8.1 represents a simple hierarchy made of six nodes and four levels. The root is located at level zero and nodes c and d are at level three. The point of this example is to identify the components of the auditory language required to undertake the sonification of this menu.

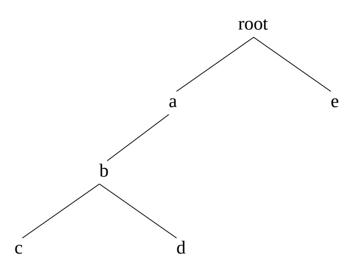


Figure 8.1: Simple menu hierarchy used in a typical sonification scenario.

8.2.1 Sounds

Imagine that the designer wants to create a sonification as described below. Let a and e be called the main menus of the hierarchy. Suppose that at this level of the hierarchy the designer would like three different sounds to convey the following pieces of information:

- A continuous sound is played as long as the user navigates within the menu branch a: this sound has to be played when the node a is selected, and keeps playing when b, c or d are selected. In addition, the intensity of this sound decreases when the depth of the current menu in the hierarchy increases: The sound has a maximum intensity at the level of a, its intensity decreases when b is selected, and decreases again if c is selected. This sound provides persistent information about which main menu branch the user is into, and information about the level. Let this sound be called Expression 1 (the first expression of the auditory language of representation being built here).

- A four second sound is played when a is selected from the root or from e, and a different four second sound is also played when e is selected from a. The former sound must be interrupted if it is not complete when a different node is selected. This sound provides information about the identity of the main menu selected by the user. Let the sound be called Expression 2.

- A very brief sound is played when a or e are selected. This sound could typically be a short percussive sound that provides key-press feedback to the user when he/she browses quickly the main menus. Let one call this sound Expression 3.

At the level below, the designer would like to carry out the following design:

- The intensity of Expression 1 is modified and a brief sound is played when b is selected. As for Expression 3, this sound can typically be a percussive sound which will be called Expression 4.

Similarly, at the level below, the designer only wants to provide the following feedback:

- Feedback to the user by playing another brief sound when c or d are selected. This sound will be called Expression 5. The control over the expressions mentioned so far can be binary (an expression can be started or stopped), or can be analogic (operated via controllers, which will be presented in the next section).

In addition, the main menu e will be sonified in a similar fashion to a. Therefore, e requires three sounds: Expression 6, Expression 7 and Expression 8, which play the same role that Expression 1, Expression 2 and Expression 3 play for a.

8.2.2 Controllers

To implement the decrease of intensity of Expression 1 described above, a controller for the volume of this expression is needed. Controller objects will be described in detail later when the global formalism will be presented. At the moment it can be considered that a controller is an object that controls the behaviour of a parameter over time. In other words, controllers provide the dynamism of the auditory expressions, and consequently of the auditor interface. In the present example, Expression 1 requires: a volume controller (Controller 1) that carries a constant value when a is selected, a similar controller (Controller 2) that carries a lower constant value when b is selected, and a third one (Controller 3) that carries a lower value when c is selected. Once a controller has been defined, it can be controlled itself by two methods: start and stop.

8.2.3 System events

It is possible to see the events of the interaction (navigation) in the menu as consisting of transitions between two menu items. In the current example, these events are represented by $(e_i)_{i=1..12}$, as shown by Figure 8.2. In this description, the list of event is not exhaustive, but is sufficient to describe the sonification intended in this example.

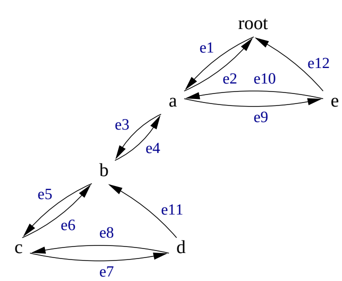


Figure 8.2: Interaction events $(e_i)_{i=1..12}$ involved in navigation in a simple menu hierarchy.

This description is sufficient to tackle the integration of sounds anywhere in the interaction process.

8.2.4 Implementation

The implementation of the sounds involves linking the interaction events with the auditory language events: each interaction event can be linked to a set of events in the auditory language. The latter events can be any of the following:

- Expression ON
- Expression OFF
- Controller ON
- Controller OFF

For the example considered in this section, the implementation can be described as follows:

- **Going down from the root to the bottom left of the hierarchy**

  This scenario involves events e_1, e_3 and e_5, sequentially:

$e_1 \longrightarrow$
  * Expression 1 ON
  * Controller 1 ON
  * Expression 2 ON
  * Expression 3 ON

$e_3 \longrightarrow$
  * Controller 1 OFF
  * Controller 2 ON
  * Expression 2 OFF
  * Expression 4 ON

$e_5 \longrightarrow$
  * Controller 2 OFF
  * Controller 3 ON
  * Expression 5 ON

- **Visiting d and coming back to c**

  This scenario involves events e_7 and e_8, sequentially:

$e_7 \longrightarrow$
  * Expression 5 ON

$e_8 \longrightarrow$ | * Expression 5 ON

- **Going back up to** a

  This scenario involves events e_6 and e_4, sequentially:

$e_6 \longrightarrow$
- * Controller 3 OFF
- * Controller 2 ON
- * Expression 4 ON

$e_4 \longrightarrow$
- * Controller 2 OFF
- * Controller 1 ON
- * Expression 3 ON

- **Selecting main menu** e

  This scenario corresponds to the event e_9:

$e_9 \longrightarrow$
- * Controller 1 OFF
- * Expression 6 ON
- * Expression 7 ON
- * Expression 8 ON

The example presented in this section has introduced the control over interaction and auditory events that is needed in menu sonification. This will be extended to a general formalism in the following section.

8.3 General Formalisation of the Integration Process

As the above example has shown, the desired sonification formalism pertains to two distinct types of objects: Auditory events (or auditory expressions) and interaction events. The formalisation of auditory expressions presented in this section was motivated by simplicity of use and implementation.

8.3.1 Expressions of the Auditory Language of Representation

The expressions of the auditory language of representation are the input of the integration process. The exact formalism of the language does not have any incidence on the sound integration process, however this formalism requires a certain form of control on the sounds to be available.

The approach chosen to formalise the auditory language of representation was to allow simple audio expressions to be created in a simple way. The aim of this research was not to design a complex musical language. Consequently, the language was based on a single simple operator allowing the creation of both sequences and chords. For convenience, a looping operator has also been added to the language. Expressions of this language can be formalised as shown below:

$$\text{Expression } e \quad \longleftarrow \left| \begin{array}{l} * \text{ ev(Parameter } p_1\text{,..., Parameter } p_m\text{)} \\ * \text{ SEQ((Expression } e_1\text{, time } t_1\text{),..., (Expression } e_n\text{, time } t_n\text{))} \\ * \text{ LOOP(Expression } e_1\text{, time } t_1\text{)} \end{array} \right.$$

Where, *ev* represents an elementary event of the language. These events can be either MIDI events or audio events *e.g.*, samples. Typically, they are controllable by a number of parameters (Parameter p_1,..., Parameter p_m) such as pitch and volume [2]. SEQ is an operator that allows expressions to be organised in sequences or chords. The syntax of the expression means that the expression e_1 starts t_1 ms after the sequence started and e_n starts t_n ms after the sequence started. LOOP is an operator that enables an expression to be played repeatedly in a loop.

For example, a simple C major chord can be represented by the following expression[3]:

SEQ(ev(60),0,ev(64),0,ev(67))

Where the events 'ev' are specified by their pitch MIDI number (for C4, E4, G4 respectively). To represent a sequence of the same notes with a duration of 500 ms, the following expression is adequate:

SEQ(ev(60,500),500,ev(64,500),500,ev(67,500),500)

[2]for MIDI events, the full list of available controllers is available at: http://www.midi.org

[3]It is assumed that the parameters of the elementary events take default values when unspecified

8.3.2 Controllers

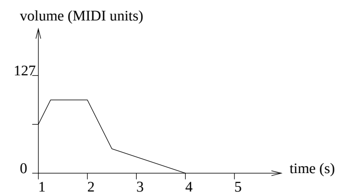


Figure 8.3: Static volume controller. The evolution of the volume is only dependent on time.

Controllers enable the parameters of a sound to be modified dynamically. For example a controller can be used to modify the volume of a sound as the sound is played. A controller is defined by an envelope that specifies the evolution of an audio parameter in function of either time or an interaction parameter. Therefore controllers can be divided in two main categories.

The first type of controllers is illustrated by Figure 8.3. The figure represents the volume envelope of a controller in function of time. When this controller is applied to a sound, the volume of the sounds follows the 4 seconds pattern represented on Figure 8.3.

On the other hand, the envelope of a controller can be a function of an interaction or system parameter. For example, in the case of a file being downloaded, a controller may be based on the amount of data downloaded so far, which is a time-dependent variable. In this context, the controller may be represented by an envelope as shown in Figure 8.4.

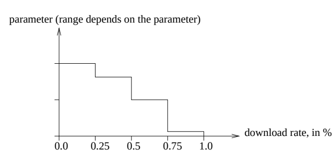


Figure 8.4: Dynamic controller. The value of the control variable is dependent on a dynamic variable of the system.

Static Controller This type of controller, illustrated in Figure 8.3, has a predetermined behaviour. It can be defined by the following parameters:

- envelope
- duration

- number of control values

Equivalently, these parameters could be used:

- envelope
- duration
- number of values per second

Or even:

- envelope
- number of values
- number of values per second

In the rest of this chapter, static controller will refer to the following definition:

$$StaticController SC = \left| \begin{array}{l} \text{Envelope E} \\ \text{Duration T} \\ \text{Number of control values N} \end{array} \right.$$

Dynamic Controller This controller, illustrated in Figure 8.4, has a behaviour determined by a dynamic variable of the system. It can be defined by the following parameters:

- Envelope
- Dynamic Variable
- Number of control values per seconds

The following dynamic constructor will be used for the remainder of this thesis:

$$DynamicController DC = \left| \begin{array}{l} \text{Envelope E} \\ \text{Dynamic Variable DV} \\ \text{Number of control values per seconds F} \end{array} \right.$$

Interaction Events

The sound interaction process would not be complete without mentioning a way to describe interaction. A lot of research has been carried out in the field of interaction modelling and a number of formalisms exist to describe various forms of interaction[4]. It would be inappropriate to propose a new formalism in the present thesis. As the AIDE architecture is modular, it would be easy to implement any interaction description formalism. The current implementation of the software is based on the simple system event description of Section 8.2.3.

8.4 Generality of the Formalism: Application to Various Sonification Cases

A common problem regarding the use of a given formalism is that it constrains designers to produce certain types of sonification which the given formalism makes easier to design. Consequently, it restrains designers' creativity. The objective of this section is to show that the formalism presented in this chapter is not biased by the primary application it is designed for: the sonification of hierarchical menus. It is important to ensure that this formalism can be applied to various sonification cases. In order to achieve this goal, two examples, unrelated to the issue of menu, sonification will be considered[5].

8.4.1 Progress-Bar Sonification I

In a recent study, Crease and Brewster [CB98] have shown that non-speech sound can be used to enhance progress bars. In their implementation of an audio progress bar, they used sound to convey several pieces of information:

- **End point sound**
  This sound indicates the target point of the task. This sound could be considered to be analogous to the right hand side of a standard graphical progress bar which fills up from left to right.

- **Progress sound**
  This sound is used to indicate the percentage of the task done. This sound could be

[4] See formalisms such as GOMS [JK96]

[5] It is not necessary to verify that the formalism applies to the sonification of hierarchical menus. The example that served as the basis for the development of the formalism ensures that this is verified.

considered to be analogous to the right hand side of the portion which is filled in a standard graphical progress bar.

- **Rate of progress sound**

  This sound is used to indicate the current rate at which the task was being completed.

- **Completion sound**

  This sound indicates that download is complete.

Walker has implemented a 3D sonification of such a progress-bar [WB99]. This section demonstrate that Walker's implementation can be formalised in the framework presented in this chapter.

For the purpose of this example, a 2D equivalent has been built. In this 2D equivalent, the download progress is represented by a "circular sound" impression: The idea is to simulate the rotation of a sound around the head of the user. When the download starts, the sound is located in front of the user's head. When 25% of the download is completed, the sound is located on the user's right side. When 50% of the download is completed, the sound is located behind the user's head. When 75% of the download is completed, the sound is located on the user's right side. This rotation occurs until the download is finished and the sound's location comes back to the front. Because the sonification only involves 2D sounds, the rotation effect is only simulated via two parameters: the panoramic position of the sound is used to represent the left-right movement of the sound, while the front-back movement is represented by the cut frequency of a filter applied to the sound. Figure 8.5 illustrates this process. On the figure, the numbers represent the download ratios. The resulting sound is represented in Sample 32.

Sample 32
Implementation of a circular auditory progress-bar. This sample represents two rotations of the download sound e.g., a download completed at a hypothetical 200%.

In order to implement this sonification, the basic sound chosen for the rotation is a filtered white noise. As figure 8.5 suggests, the rotation effect is achieved by modifying the cut frequency of the filter and the panoramic position of the sound. The rest of this section describes how this sonification can be implemented with the formalism described above.

The sonification involves one expression: exp, a continuous noise of indefinite duration:

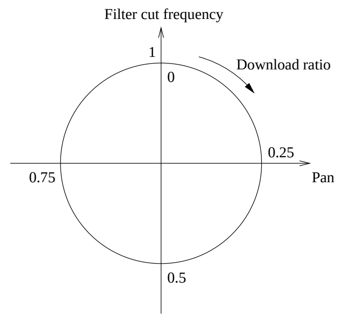


Figure 8.5: Circular progress-bar. The two parameters (x and y axes) of the rotation are panoramic and a filter cut frequency.

$exp \leftarrow noise$

In addition, two controllers are required to implement the rotation represented on Figure 8.5: DC1, defined below, represents the panoramic location of the sound (horizontal axis on Figure 8.5). This controller uses an envelope (env_1) represented in Figure 8.6. Similarly, DC2, also defined below, represents the filter cut frequency (vertical axis on Figure 8.5). This controller uses an envelope (env_2) represented in Figure 8.7. As Figures 8.6 and 8.7 show, the perception of a circular motion can be approximated by using linear envelopes instead of sine waves.

$DC1 \leftarrow DC(env_1, e_3, 10)$[6]

$DC2 \leftarrow DC(env_2, e_3, 10)$

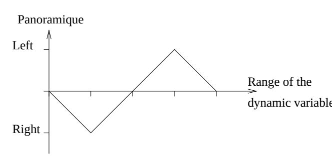


Figure 8.6: env_1, envelope of controller $DC1$.

As far as the system is concerned, two static events and one dynamic events are necessary to describe the download. The two static events, e_1 and e_2 respectively represent the beginning and the end of the download. The dynamic event e_3 represents the download ratio (between 0

[6] e_3 is a dynamic system event defined later. It represents the download rate.

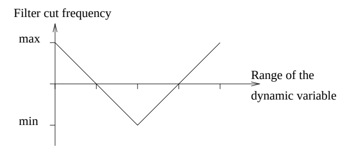


Figure 8.7: env_2, envelope of controller $DC2$.

and 1) and is used by the audio controllers $DC1$ and $DC2$.

Now that the elements of the sonification have been described (both the expressions of the auditory language of representation and the system/interaction events), the sonification can be formalised very easily. In the following expressions, the static system events are simply associated with at least one audio expression or controller:

$$e_1 \longrightarrow \left| \begin{array}{l} * \text{ play } exp \\ * \text{ start } DC1(exp) \\ * \text{ start } DC2(exp) \end{array} \right.$$

$$e_2 \longrightarrow \left| \; * \text{ stop } exp \right.$$

In order to demonstrate the flexibility of the formalism, the next example extends the sonification described above.

8.4.2 Progress-Bar Sonification II

The sonification proposed here is a simple extension of the previous one. The idea behind this extended sonification is to convey information at different stages of the download rather than continuously. This can be done by using the previous example, and only playing the sound when the download starts, when 25%, 50%, 75% of the file has been downloaded, and finally when the download ends (see Sample 33).

Sample 33

Implementation of a circular audio progress-bar that notifies progress periodically.

The only modification to the sonification described above involves using an additional volume controller to the audio expression *exp*. The envelope used by this controller is displayed in Figure 8.8. The extended sonification is formalised below:

$$e_1 \longrightarrow \left| \begin{array}{l} * \text{ play } exp \\ * \text{ start } DC1(exp) \\ * \text{ start } DC2(exp) \\ * \text{ start } DC3(exp) \end{array} \right.$$

Where $DC3$ is defined by:

$$DC3 \leftarrow DC(env_3, e_3, 10)$$

$$e_2 \longrightarrow \left| \ * \text{ stop } exp \right.$$

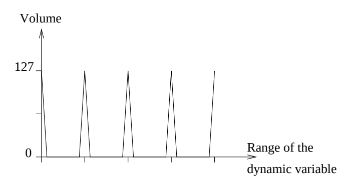


Figure 8.8: env_3, envelope of the dynamic controller $DC3$.

The three examples examined above show that the sonification formalism defined in this chapter in not only appropriate to the sonification of hierarchical menus. It is also appropriate to completely different sonification cases. In fact, it can reasonably be extrapolated that this formalism is general enough to support most sonification problems.

8.5 Implementation of the Formalism in AIDE

So far, this chapter has introduced the need for appropriate tools to carry out the implementation of sounds in interface prototypes. Then a general sonification formalism has been presented and its generality has been reviewed. The current section describes a tool based on this formalism: AIDE (Auditory Interface Design Environment).

The development of AIDE was in fact motivated by the lack of flexibility of the various tools (sequencers especially) used while designing large sets of hierarchical sounds for mobile telephone menus: In Chapter 4 there were three stages in the design and implementation of the sounds in the telephone menu. Firstly, the sounds were designed with a sequencer (Cakewalk); then they were recorded one by one as audio files with Steinberg Wavelab; finally, they were attached to each node of the hierarchy. This process implied that if a modification had to be made to a sound, it would have had to be recorded again. Moreover, because the sounds were linked to one another by inheritance relationships, modifying one sound would most likely mean the modification and re-recording of a large part of the sonification. Carrying out the time consuming development of a software tool was therefore not only the consequence of the theoretical reflection on sound design, but was practically needed to pursue this research further.

The approach to the development of AIDE was to develop an environment that would meet two requirements:

1. Allow further audio design and implementation to be carried out with AIDE rapidly.
2. provide an architecture flexible enough to support further development.

The latter requirement was achieved by carrying out a modular implementation of the software according to which each module can be improved independently from each other. The former requirement involved a list of more specific important requirements:

Visual Design A graphical user interface is required for the creation and manipulation of auditory objects.

Easy connection between interaction events and auditory events The interface should allow easy integration and modification of the sounds designed with the tool.

Shortcut to carry out multiple connections between auditory events and interaction events In a large telephone menu hierarchy, one sound may be allocated to many different menus. To facilitate this operation, AIDE provides adequate functions.

Possibility to create auditory expressions from existing expressions This enables hierarchical links between expressions to be created. Subsequently, when an expression e is modified, the modification is propagated to all expression in which e appears.

Rapid prototyping capabilities When a sonification is being designed, a prototype of the sonified interface should be available at any time of the design process.

The next section provides details for each of these requirements.

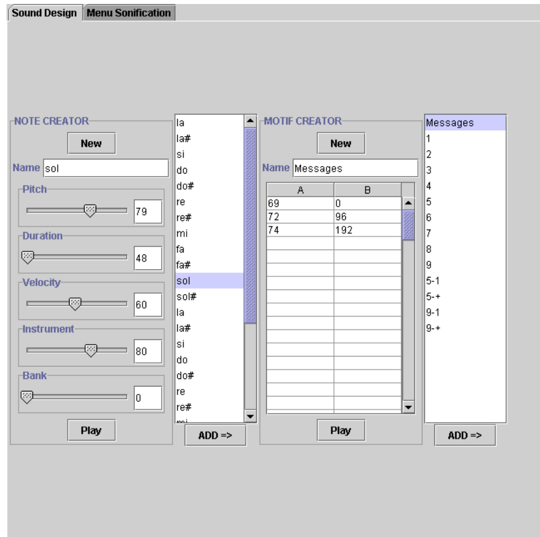


Figure 8.9: Graphical User Interface for the design of auditory expressions.

8.5.1 A Visual Design Environment for Auditory Interfaces

Visual environments are proven to be a very effective way of designing interfaces. To create music or sounds, graphical user interfaces are also indispensable. In AIDE, the creation and manipulation of sounds and their integration in an interface prototype is entirely supported by a graphical user interface. Figure 8.9 shows the GUI for the creation of basic motifs. The left hand panel ("Note Creator") is used to create notes. A number of parameters of these notes can be adjusted. These parameters are the standard MIDI parameters: pitch, duration, velocity instrument and bank. The parameters can be adjusted with a slider or values can be entered directly in the text box. The note creation panel also features two buttons. One to create a new note, which is added to the list of notes available, and one to play the current note. Next to the note creation panel is the list of available notes. Notes are added to this list when they are created. This list allows standard editing operations to be carried out, such as selecting single or multiple notes and deleting them. Selected notes can also be added to a motif. When this operation is carried out, the list of notes are added to the "Motif Creator" table. This table contains a list of notes and the starting time of these notes in milliseconds. On figure 8.9, the motif represented in the motif creation panel is called "Messages", it contains three notes named after their pitch starting at times: 0, 96 and 192 milliseconds. All the motifs created are listed in a list displayed next to the motif creation panel.

Once a motif has been created, it can be added to a list of auditory expressions that can be integrated in an interface prototype[7]. Figure 8.10 shows the panel which is the interface between the sound design part of the tool and the prototyping part of the tool. This panel also allows the designer to perform various transformations of the motifs such as, transposition, time stretching, modification of the volume and so on. It is also possible to create new expressions from sequences of other expressions.

All expressions created or modified can be selected in this panel (the list of selected expressions is displayed at the bottom of the panel on Figure 8.10) and subsequently associated to the prototype (in this case the nodes of the menu). This association between nodes and sounds can be accomplished from the panel represented on Figure 8.11 or from the *Sonification* menu of the application. Figure 8.11 shows that an auditory event can be associated to any of four interaction events related to a node. These events represent the means by which a node can be selected: from its ancestor node (by pressing the select key on the mobile phone), from each of the two nodes at the same level (up and down arrow keys), or from a node at the level below (back key). Once a sound has been allocated to a node, the node is displayed in blue in the tree representation of the menu.

Finally in AIDE, the mobile phone simulation panel allows the designer to instantly try the sonification he/she is working on. The simulation represented on Figure 8.12 reproduces the behaviour of the NOKIA *6110* menu.

The way the panels described so far are laid out in the AIDE GUI is based on the sonification process flow. On the right of the application, the interface panel (between the sound design part and the prototyping part) is always visible. The larger left panel displays either the sound design tabbed pane or the menu representation and phone simulation tabbed pane. Figure 8.13 represents the whole GUI, when the prototyping pane is selected. As the picture shows, the tool can be used in two modes, depending which pane is selected. When the “Sound Design” pane is selected, sounds can be created and added to a container located on the right hand side of the tool GUI. When the “Menu sonification” pane is selected (as shown in Figure 8.13), the sounds listed in the container can be selected and added to nodes of the hierarchy.

[7]The interface prototype part is limited to mobile phone menus in the version of the software presented in this thesis

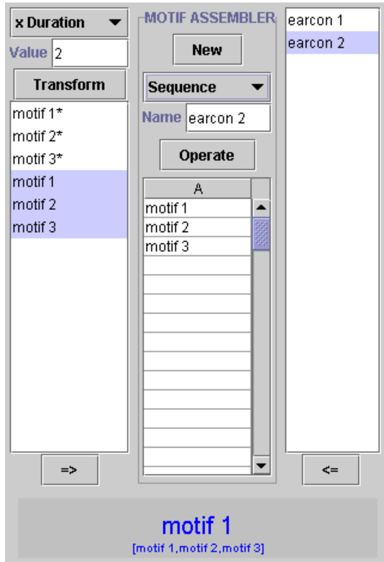


Figure 8.10: Interface between the sound design part of the software and the prototyping part of the software.

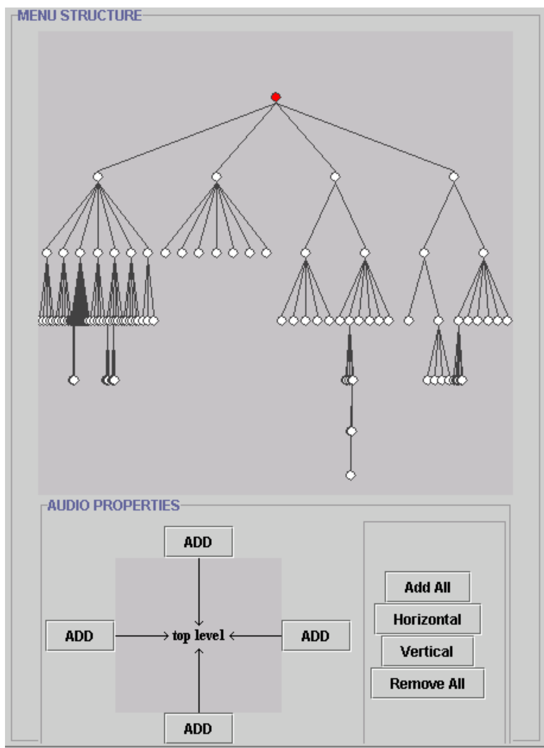


Figure 8.11: Panel representing the structure of this interface and the interaction events (*e.g.*, the menu navigation events).

Figure 8.12: Panel representing the mobile phone menu prototype.

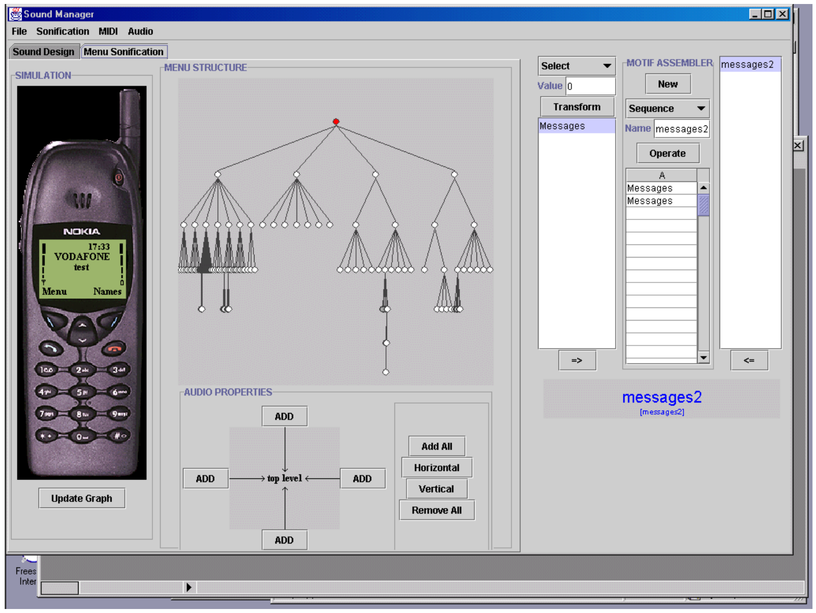


Figure 8.13: Overall view of the AIDE interface.

8.5.2 Easy Connection Between the Interaction and the Sounds

The integration of sounds with the hierarchical menu can be carried out easily and in different ways. Firstly, if an auditory expression is selected (it is then displayed in blue at the bottom of the panel represented in Figure 8.10), it can be associated individually to any of the four interaction events of a node, or to all four (with the *add all* button). An audio expression can be disconnected from one or more interaction event of a node as easily. These operations can be performed from the *sonification* menu as well. The *sonification* menu also contains functionality that speeds up the sonification process: a sound or a list of sounds can be connected to all the nodes at a given level of a branch in the hierarchy with a single operation. It is also possible to disconnect the sounds from a group of nodes in a similar fashion.

8.5.3 Management of Structural Relationships Between Auditory Expressions

One of the main reasons why the development of AIDE was undertaken was the lack of possibilities provided by other tools to manage the structural relationships between sounds once integrated in the menu. AIDE has been designed so that structural relationships between auditory expressions are easy to manage. Here is an example that demonstrates this point: if a set of auditory expressions B_1, B_2 and B_3 are created from an expression A, and if A is modified, B_1, B_2 and B_3 will be modified accordingly. This novel feature constitutes an important part of the flexibility of AIDE and is not provided by any other prototyping tool.

8.5.4 Rapid Prototyping

In human-computer interface design, being able to test the interface during the design constitutes an important asset. AIDE provides this facility by allowing the designer to navigate in the menu at any point of the design process.

8.6 Conclusions and Future Developments

The auditory language presented at the beginning of the current chapter was partially implemented in the version of AIDE described in the previous sections. As mentioned earlier, the focus of the present thesis is not the development of a sound design tool. AIDE was developed to facilitate the design and evaluation of large sets of hierarchical sets of sounds and to provide a framework for the extension of the tool to other sonification areas.

One of the strength of AIDE as far as future developments are concerned is its architecture. AIDE involves two distinct modules, sound module and interaction module, which can be extended or modified independently. Future developments have already occurred with two MSc students working on the improvement of each module. On the one hand, one project focused on providing a better GUI for the sound design module. The outcome of this project was a sequencer-like GUI that allows the designer to create audio motifs more efficiently. The full implementation of the auditory language described earlier in this chapter remains to be done. On the other hand the development of the interaction module focused on two points addressed by the second MSc project. The first part of the project aims at allowing AIDE to support the sonification of other kinds of interfaces than mobile telephone menus. In particular, it should support the sonification of widgets and groups of widgets. The second more challenging objective of this project was to extend the notion of interaction events related to menu navigation currently implemented in AIDE. In the current framework, an interaction event corresponds to the selection of a menu item. The four possible interaction events for a given item involve the four possible directions from which the selection can be operated. Although this framework enables the implementation of effective sonifications, the conceptualisation of the auditory interaction would benefit from a richer description of the interaction.

Chapter 9

Conclusion

Motivated by the evidence of interaction issues related to navigation in complex hierarchical structures in interfaces with restricted display, this thesis has studied the use of non-speech audio to address these issues. This chapter summarises the contribution of this thesis, highlights the strengths and weaknesses of this research and justifies the need for further work.

9.1 Summary of the Contribution

The chief contribution of this thesis pertains to the use of non-speech sounds in mobile devices such as mobile telephones:

- A framework for the design of sounds to support navigation in hierarchical menus has been defined, based on a careful study of navigation. Design principles and implementation guidelines have been devised from this theoretical study.
- This framework has been put to practice in two instances, for the sonification of two large mobile telephone menus. These applications have demonstrated that the design principles can be applied in very distinct contexts.
- The sonification designed in the ideal case has been evaluated and demonstrated that our design framework led to an improvement of the usability of a mobile telephone menu.
- The sonification process which was at the centre of this research has been formalised, and provided a solid basis for the development of a sonification tool. A first version of such a tool has been developed and used for one of our designs.

9.2 Originality and Strengths

The approach followed to tackle the issue of improving navigation with non-speech sounds is the main originality of this thesis. The research was motivated by the following question: how to use the audio channel to improve navigation in interfaces with restricted display?

Some Fundamental Principles

One of the strengths of this research was to focus on navigation rather than on menu structure, in order to devise a design framework. In a menu, the sounds should not only be mapped to the syntactic structure of the menu, but also its semantic structure. In addition, sound should also provide information that supports the motion of the user in the menu, the same way this motion is supported in desktop menus.

Implementation Recommendations

The advantage of high-level principles is that they apply to many different cases. The drawback however is that it can be difficult for a designer to apply these general principles to a particular case. Therefore some implementation recommendations are useful to guide designers through the implementation of sonifications in the most common cases.

The recommendations proposed in Chapter 5 assist designers on two fronts:

- Understanding in which case to apply the principles, in function of the menu properties. For example it has been recommended to apply Principle 1 to design sounds for the main menu options.
- Using the appropriate type of sounds while implementing these principles. The only limit to the application of the design principles is the designer's imagination. It is also the designer's knowledge of sound psychology. The design guidelines of Chapter 5 proposed various approaches to the design of sound coherent with sound psychology results reviewed in Chapter 3.4.

Applications

This thesis has demonstrated in practice that the principles can be applied in two very different contexts with regards to the constraints imposed on sound design. The first application has shown that even when the implementation guidelines did not apply to the context, it was possible to refer to the high-level design principles to devise a complex menu sonification for a real mobile telephone prototype. The second application has demonstrated that with no constraint on sound design, it was possible to devise a completely different sonification for a similar menu.

Evaluation

The strengths of the evaluation carried out in Chapter 7 are the following:

- It involved a real mobile telephone menu.
- The sonified menu investigated was much larger than the menus investigated in previous research [Bre98].
- The evaluation was focusing on the performance of actual navigation tasks, not on the understanding or recollection of the sonification.
- Despite the fact that there was only one iteration to the design of the sonification, and that it was the first time a menu sonification of this scale was carried out (166 different sounds involved), the experiment showed that the sonification had a very positive effect on the performance of the tasks investigated.

AIDE

The design carried out in Chapter 7 made it painfully clear that there was no appropriate tools available to undertake a sonification of that type and scale. Chapter 8 has addressed this issue by investigating the formalism required for the development of an auditory interface prototyping tool. This formalism involves an auditory language of representation, that allows designers to create sounds that can be linked to each other. In addition, the formalism enables the sounds to be easily connected to an interface.

The soundness of the formalism has been demonstrated by showing that it does not only apply to the sonification of menus, but also to a totally different sonification problem (sonification of a progress bar). In addition, the most obvious demonstration of the pertinence of the formalism was its implementation in AIDE, an Auditory Interface Design Environment, which was then used to carry out the sonification of a mobile telephone prototype menu (see Chapter 6). During the course of this sonification, AIDE proved to be far more flexible design solution than using a sequencer such as Cakewalk.

9.3 Limits and Perspectives

The main weakness of this research is linked to the novelty of the design framework presented. In Chapter 5, the implementation guidelines do not cover all possible design contexts, and the two applications proposed (in Chapters 6 and 7) only cover two. Therefore, further implementation would be a valuable complement to the work presented in this thesis. Further design would enable more cases to be tackled, but may also improve the current deisigns. For instance, some negative subjective feedback has been collected in the experiments presented in Chapters 7 and 6. If more time had been allocated to carry out additional research, the design of the sounds used could have been improved based on the user's feedback.

This thesis would also benefit from exhibiting further concrete applications of the fifth design principle: providing users with various levels of sonifications. Because sound is intrusive by nature and because users tolerate sounds in a very variable way, it is important for the density of a sonification to be adjustable. In addition, because users' tastes in music and regarding sound qualities also vary a lot, the aesthetic qualities of a sonification should also be adjustable. Therefore, the configurability and customisation of auditory interfaces is an issue would a very useful application of the framework proposed in this thesis.

In addition, because the principles and the design approach are new, they have only been applied by the author. Additional designs by other sound designers would constitute a valuable extension to the research completed during the course of this thesis.

The auditory interface design environment described in Chapter 8 would benefit from further development. The state of the development of AIDE is currently restricted to the sonification of hierarchical menus. However the architecture of the software is flexible enough to enable further development: the sound creation part of the software is independent from the prototyping part of the software. Therefore, they can both be improved independently. In fact, two Master's projects at the University of Glasgow have already addressed these two aspects independently. The first project has looked at the improvement of the sound creation part of the tool. As a result of this project, the graphical interface for this part of the tool was transformed into a sequencer-like interface, making the creation of auditory motifs easier. The second Master's project has investigated ways to make the user interface prototyping way more generic. Currently, this part of the tool only support hierarchical menus. With this Master's work, the tool allows the user to describe the interaction they want to support with sound.

AIDE is currently targetted at sound designers. The solution to make it accessible to the HCI community in general involves pursuing the investigation in automatic sound generation proposed in this thesis. Because automatic sound generation involves a combination of non-trivial Artificial Intelligence and a thorough understanding of sound design, more research is required to achieve the automatic generation of complex usable sonifications.

In addition to further sound design and software development, more resources or further work would have helped understanding more about navigation patterns in large menu hierarchies. Menu navigation is in itself a vast research field which was not the focus of this thesis, but whose understanding was fundamental in devising sound design principles. During the experiment presented in Chapter 7, exhaustive data concerning participants' navigation were collected. For the purpose of the analysis carried out in that chapter, only quantitative information were analysed. However it would have been interesting to investigate qualitative properties of the participants' performances: Were they searching deep in the tree, were they browsing the branches exhaustively, what kind of path were they following?

Applications

One of the strengths of the principles proposed in this thesis is that they are general enough to be applied to the sonification of various types of interfaces, from telephone-based spoken menus, to WAP menus. Unfortunately, only a couple of applications can be treated in one thesis, and hence only the sonification of mobile telephone menus has been addressed in details (in Chapters 6 and 7).

The application of our principles to menus in different types of interface would constitute further evidence of their generality and pertinence. In particular, it would be interesting to apply them to very distinct types of interfaces: on the one hand telephone-based services, referred to as the first category of TBIs in Chapter 2, whose output is auditory only; on the other hand, PDAs, which offer a richer form of interaction with multimedia content.

Devices with a rich interaction bandwidth are a potentially fruitful area of applications of the results presented in this thesis. Conversly, purely auditory interfaces would also be suited to applying these results. It has been proven that blind users can benefit from the use of non-speech sounds in auditory interfaces. With a little more research, the principles developed in this thesis could be combined with spoken menu design principles to enable blind users to navigate in large hierarchical menus. Furthermore, the results of the experiment presented in Chapter 7 suggest that, with a little bit of experience, non-sighted users could navigate in menu hierarchies quickly, relying on non-speech sounds only.

9.4 Final Word

With the design principles, large-scale menu sonifications and evaluations, sound design formalism and tool presented in this thesis, this research constitutes a valuable contribution to the knowledge and investigation of the many issues involves in supporting interaction with non-speech audio. In particular, this thesis should provide a valuable resource to tackle the design of sounds for limited display devices and pursue research in this area.

Appendix A

Experiment 1 Raw Data

The following tables present the data collected in the experiment reported in chapter 4. The first four tables contain the data collected for each family for the musician group. The following four tables contain the data collected for the same four families with the non-musician group. Finally the last two tables present the overall results for both groups.

In the following eight tables, each row represents the recall scores of a subject. Each table contains eleven rows that correspond to the eleven subjects of each group. Each row contains four sub-rows labelled ok, l, f, sf and t. The value of ok (0 or 1) represents whether the node is recalled correctly (1) or not (0). When a node has not been recalled correctly, l, f and sf represent the component of the sound that has not been recalled correctly. When a node has not been recalled properly, l represents the value of the wrong node, f represents its family, and sf represents its sub-family. In addition t represents the number of components recalled properly. If a node is recalled accurately, t takes the value of three.

		1.1	1.1.1	1.1.2	1.1.3	1.1.1.1	1.1.2.1
I	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
II	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
III	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
IV	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
V	ok	1	0	1	1	0	0
	l	0	0	0	0	0	0
	f	0	4	0	0	0	2
	sf	0	0	3	0	2	0
		3	2	3	3	2	2
VI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VIII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
IX	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
X	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
XI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
	marks	33	32	33	33	32	32
	%	100.00	96.97	100.00	100.00	96.97	96.97
	total	**11.00**	**10.00**	**11.00**	**11.00**	**10.00**	**10.00**
	%	**100.00**	**90.91**	**100.00**	**100.00**	**90.91**	**90.91**

Musician scores – family 1.

		1.2	1.2.1	1.2.2	1.2.3	1.2.1.1	1.2.2.1
I	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
II	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
III	ok	1	1	1	0	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	2	0	0
		3	3	3	2	3	3
IV	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
V	ok	0	0	0	1	1	0
	l	0	0	4	0	0	0
	f	3	4	0	0	0	0
	sf	0	0	0	0	0	3
		2	2	2	3	3	2
VI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VIII	ok	1	1	1	1	1	0
	l	0	0	0	0	0	3
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	2
IX	ok	0	1	1	1	1	1
	l	0	0	0	0	0	0
	f	3	0	0	0	0	0
	sf	0	0	0	0	0	0
		2	3	3	3	3	3
X	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
XI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
	marks	31	32	32	32	33	31
	%	93.94	96.97	96.97	96.97	100.00	93.94
	total	**9.00**	**10.00**	**10.00**	**10.00**	**11.00**	**9.00**
	%	**81.82**	**90.91**	**90.91**	**90.91**	**100.00**	**81.82**

Musician scores – family 2.

		1.3	1.3.1	1.3.2	1.3.3	1.3.1.1	1.3.3.1
I	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
II	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
III	ok	1	0	1	1	1	1
	l	0	0	0	0	0	0
	f	0	3	0	0	0	0
	sf	0	0	0	0	0	0
		3	2	3	3	3	3
IV	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
V	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VII	ok	1	1	1	1	0	1
	l	0	0	0	0	0	0
	f	0	0	0	0	4	0
	sf	0	0	0	0	0	0
		3	3	3	3	2	3
VIII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
IX	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
X	ok	1	1	1	1	0	1
	l	0	0	0	0	3	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	2	3
XI	ok	1	1	1	1	0	1
	l	0	0	0	0	3	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	2	3
	marks	33	32	33	33	30	33
	%	100.00	96.97	100.00	100.00	90.91	100.00
	total	**11.00**	**10.00**	**11.00**	**11.00**	**8.00**	**11.00**
	%	**100.00**	**90.91**	**100.00**	**100.00**	**72.73**	**100.00**

Musician scores – family 3.

		1.4	1.4.1	1.4.2	1.4.3	1.4.2.1	1.4.3.1
I	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
II	ok	1	1	1	1	0	0
	l	0	0	0	0	3	0
	f	0	0	0	0	3	3
	sf	0	0	0	0	0	0
		3	3	3	3	2	2
III	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
IV	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
V	ok	0	1	1	0	0	1
	l	4	0	0	4	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	3	0
		2	3	3	2	2	3
VI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
VIII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
IX	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
X	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
XI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
	marks	32	33	33	32	31	32
	%	96.97	100.00	100.00	96.97	93.94	96.97
	total	**10.00**	**11.00**	**11.00**	**10.00**	**9.00**	**10.00**
	%	**90.91**	**100.00**	**100.00**	**90.91**	**81.82**	**90.91**

Musician scores – family 4.

		1.1	1.1.1	1.1.2	1.1.3	1.1.1.1	1.1.2.1
I	ok	1	0	1	1	1	1
	l	0	2	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	2	3	3	3	3
II	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
III	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
IV	ok	1	1	1	1	0	1
	l	0	0	0	0	0	0
	f	0	0	0	0	2	0
	sf	0	0	0	0	0	0
		3	3	3	3	2	3
V	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
VI	ok	1	1	0	1	1	1
	l	0	0	3	0	0	0
	f	0	0	3	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	1	3	3	3
VII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
VIII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
IX	ok	1	1	0	0	1	1
	l	0	0	0	4	0	0
	f	0	0	3	3	0	0
	sf	0	0	0	0	0	0
	t	3	3	2	1	3	3
X	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
XI	ok	1	1	1	1	0	0
	l	0	0	0	0	0	0
	f	0	0	0	0	0	3
	sf	0	0	0	0	2	3
	t	3	3	3	3	2	1
	marks	33	32	30	31	31	31
	%	100	96.9697	90.90909	93.93939	93.93939	93.93939
	total	**11**	**10**	**9**	**10**	**9**	**10**
	%	**100**	**90.90909**	**81.81818**	**90.90909**	**81.81818**	**90.90909**

Non-musician scores – family 1.

		1.2	1.2.1	1.2.2	1.2.3	1.2.1.1	1.2.2.1
I	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
II	ok	0	1	0	1	1	0
	l	0	0	0	0	0	3
	f	3	0	0	0	0	0
	sf	0	0	1	0	0	3
	t	2	3	2	3	3	1
III	ok	1	1	1	0	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	2	0	0
	t	3	3	3	2	3	3
IV	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
V	ok	1	1	1	0	0	0
	l	0	0	0	4	0	0
	f	0	0	0	0	0	4
	sf	0	0	0	2	2	0
	t	3	3	3	1	2	2
VI	ok	1	0	1	0	0	0
	l	0	0	0	0	0	3
	f	0	4	0	3	0	3
	sf	0	0	0	2	2	0
	t	3	2	3	1	2	1
VII	ok	1	1	1	1	1	0
	l	0	0	0	0	0	0
	f	0	0	0	0	0	3
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	2
VIII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
IX	ok	1	1	1	1	1	0
	l	0	0	0	0	0	0
	f	0	0	0	0	0	4
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	2
X	ok	1	1	1	0	1	1
	l	0	0	0	3	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	2	0	0
	t	3	3	3	2	3	3
XI	ok	1	1	1	0	0	1
	l	0	0	0	2	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	2	0
	t	3	3	3	2	2	3
	marks	32	32	32	26	30	26
	%	96.9697	96.9697	96.9697	78.78788	90.90909	78.78788
	total	**10**	**10**	**10**	**7**	**9**	**6**
	%	**90.90909**	**90.90909**	**90.90909**	**63.63636**	**81.81818**	**54.54545**

Non-musician scores – family 2.

		1.3	1.3.1	1.3.2	1.3.3	1.3.1.1	1.3.3.1
I	ok	1	1	1	0	1	1
	l	0	0	0	0	0	0
	f	0	0	0	2	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	2	3	3
II	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
III	ok	1	0	1	1	0	1
	l	0	0	0	0	0	0
	f	0	3	0	0	4	0
	sf	0	0	0	0	2	0
	t	3	2	3	3	1	3
IV	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
		3	3	3	3	3	3
V	ok	0	1	1	0	0	1
	l	0	0	0	0	3	0
	f	2	0	0	2	0	0
	sf	0	0	0	0	0	0
	t	2	3	3	2	2	3
VI	ok	1	1	1	1	0	1
	l	0	0	0	0	3	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	2	3
VII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
VIII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
IX	ok	1	0	1	1	1	1
	l	0	0	0	0	0	0
	f	0	2	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	2	3	3	3	3
X	ok	1	1	1	1	0	1
	l	0	0	0	0	0	0
	f	0	0	0	0	1	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	2	3
XI	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
	marks	32	31	33	31	28	33
	%	96.9697	93.93939	100	93.93939	84.84848	100
	total	**10**	**9**	**11**	**9**	**7**	**11**
	%	**90.90909**	**81.81818**	**100**	**81.81818**	**63.63636**	**100**

Non-musician scores – family 3.

		1.4	1.4.1	1.4.2	1.4.3	1.4.2.1	1.4.3.1
I	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
II	ok	1	1	1	1	0	0
	l	0	0	0	0	3	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	2
	t	3	3	3	3	2	1
III	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
IV	ok	0	1	0	0	0	0
	l	0	0	4	4	3	0
	f	3	0	0	0	0	0
	sf	0	0	0	0	0	2
		2	3	2	2	2	2
V	ok	0	0	0	0	0	0
	l	3	4	4	4	0	0
	f	0	0	0	0	3	3
	sf	0	2	0	2	1	0
	t	2	1	2	1	1	2
VI	ok	0	0	1	0	1	1
	l	3	4	0	4	0	0
	f	2	0	0	2	0	0
	sf	0	3	0	2	0	0
	t	1	1	3	0	3	3
VII	ok	0	1	1	1	1	1
	l	3	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	2	3	3	3	3	3
VIII	ok	1	1	1	1	1	1
	l	0	0	0	0	0	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	3	3
IX	ok	1	1	0	1	0	0
	l	0	0	0	0	0	0
	f	0	0	4	0	2	3
	sf	0	0	0	0	0	0
	t	3	3	2	3	2	2
X	ok	1	1	1	1	0	1
	l	0	0	0	0	3	0
	f	0	0	0	0	0	0
	sf	0	0	0	0	0	0
	t	3	3	3	3	2	3
XI	ok	0	0	0	0	0	0
	l	0	0	0	0	3	0
	f	0	0	0	0	2	0
	sf	3	3	3	1	3	2
	t	2	2	2	2	0	2
	marks	27	28	29	26	24	27
	%	81.81818	84.84848	87.87879	78.78788	72.72727	81.81818
	total	**6**	**9**	**7**	**8**	**5**	**6**
	%	**54.54545**	**81.81818**	**63.63636**	**72.72727**	**45.45455**	**54.54545**

Non-musician scores – family 4.

	marks	%	total	%	subtotals
1.1	33	100	11	100	
1.1.1	32	96.9697	10	90.90909	
1.1.2	33	100	11	100	
1.1.3	33	100	11	100	
1.1.1.1	32	96.9697	10	90.90909	
1.1.2.1	32	96.9697	10	90.90909	95.45455
1.2	31	93.93939	9	81.81818	
1.2.1	32	96.9697	10	90.90909	
1.2.2	32	96.9697	10	90.90909	
1.2.3	32	96.9697	10	90.90909	
1.2.1.1	33	100	11	100	
1.2.2.1	31	93.93939	9	81.81818	89.39394
1.3	33	100	11	100	
1.3.1	32	96.9697	10	90.90909	
1.3.2	33	100	11	100	
1.3.3	33	100	11	100	
1.3.1.1	30	90.90909	8	72.72727	
1.3.3.1	33	100	11	100	93.93939
1.4	32	96.9697	10	90.90909	
1.4.1	33	100	11	100	
1.4.2	33	100	11	100	
1.4.3	32	96.9697	10	90.90909	
1.4.2.1	31	93.93939	9	81.81818	
1.4.3.1	32	96.9697	10	90.90909	92.42424
total	773	97.60101	245	92.80303	
%			92.80303		

Musicians overall scores.

	marks	%	total	%	subtotals
1.1	33	100	11	100	
1.1.1	32	96.9697	10	90.90909	
1.1.2	30	90.90909	9	81.81818	
1.1.3	31	93.93939	10	90.90909	
1.1.1.1	31	93.93939	9	81.81818	
1.1.2.1	31	93.93939	10	90.90909	89.39393939
1.2	32	96.9697	10	90.90909	
1.2.1	32	96.9697	10	90.90909	
1.2.2	32	96.9697	10	90.90909	
1.2.3	26	78.78788	7	63.63636	
1.2.1.1	30	90.90909	9	81.81818	
1.2.2.1	26	78.78788	6	54.54545	78.78787879
1.3	32	96.9697	10	90.90909	
1.3.1	31	93.93939	9	81.81818	
1.3.2	33	100	11	100	
1.3.3	31	93.93939	9	81.81818	
1.3.1.1	28	84.84848	7	63.63636	
1.3.3.1	33	100	11	100	86.36363636
1.4	27	81.81818	6	54.54545	
1.4.1	28	84.84848	9	81.81818	
1.4.2	29	87.87879	7	63.63636	
1.4.3	26	78.78788	8	72.72727	
1.4.2.1	24	72.72727	5	45.45455	
1.4.3.1	27	81.81818	6	54.54545	62.12121212
total	715	90.27778	209	79.16667	
%			77.65152		

Non-musicians overall scores.

Appendix B

Detailed Structure of the NOKIA *8210* Menu

B.1 Presentation

This appendix presents the complete structure of the NOKIA *8210* menu used in the sonification designed in Chapter 6. The structure is presented as a flat textual representation. This format is the same as that used by AIDE for the description of hierarchical menu (See Chapter 8 and Appendix D).

In the following section, each line represents a menu option. The first field of the line represents the number of submenus of the current option and the submenus are listed immediatly after their parent. The second field represents the type of menu option while the third field, when it exists, is the label of the menu option. More information on the format of these menus is available in the AIDE manual provided on the attached CD (See Appendix D).

B.2 Menu Structure

11,start 9,one,Messages, , , ,0,1,1,Select,Exit
0,one,Inbox, , , ,0,1,0,Select,Back
0,one,Outbox, , , ,0,1,0,Select,Back
0,one,Write,messages, , ,0,1,0,Select,Back
0,one,Picture,messages, , ,0,1,0,Select,Back
3,one,Message,settings, , ,0,1,0,Select,Back
4,one,Set 1, , , ,0,1,0,Select,Back
0,one,Message center,number, ,

,0,1,0,Select,Back

0,one,Messages,sent as, , ,0,1,0,Select,Back

0,one,Message,validity, , ,0,1,0,Select,Back

0,one,Rename,this set, , ,0,1,0,Select,Back

4,one,Set 2, , , ,0,1,0,Select,Back

0,one,Message center,number, ,

,0,1,0,Select,Back

0,one,Messages,sent as, , ,0,1,0,Select,Back

0,one,Message,validity, , ,0,1,0,Select,Back

0,one,Rename,this set, , ,0,1,0,Select,Back

2,one,Common, , , ,0,1,0,Select,Back

0,one,Delivery,report, , ,0,1,0,Select,Back

0,one,Reply via,same centre, ,

,0,1,0,Select,Back

0,one,Info,service, , ,0,1,0,Select,Back

0,one,Fax or,data call, , ,0,1,0,Select,Back

0,one,Service,command,editor,

,0,1,0,Select,Back

2,one,Voice,messages, , ,0,1,0,Select,Back

0,one,Listen to,voice messages, ,

,0,1,0,Select,Back

0,one,Voive mailbox,number, ,

,0,1,0,Select,Back

7,one,Call Register, , , ,0,1,1,Select,Exit

0,one,Missed calls, , , ,0,1,0,Select,Back

0,one,Received calls, , , ,0,1,0,Select,Back

0,one,Dialled,numbers, , ,0,1,0,Select,Back

0,one,Erase recent,call lists, ,

,0,1,0,Select,Back

5,one,Show call,duration, , ,0,1,0,Select,Back

0,one,Last call,duration, , ,0,1,0,Select,Back

0,one,All calls',duration, , ,0,1,0,Select,Back

0,one,Received calls',duration, ,

,0,1,0,Select,Back

0,one,Dialled calls,duration, ,

,0,1,0,Select,Back

0,one,Clear,timers, , ,0,1,0,Select,Exit

3,one,Show call,costs, , ,0,1,0,Select,Back

0,one,Last call,cost, , ,0,1,0,Select,Back

0,one,All calls',cost, , ,0,1,0,Select,Back

0,one,Clear,counters, , ,0,1,0,Select,Back

2,one,Call costs',settings, , ,0,1,0,Select,Back

0,one,Call costs',limit, , ,0,1,0,Select,Back

0,one,Show,costs in, , ,0,1,0,Select,Back

5,one,Profiles, , , ,0,1,1,Select,Exit

2,all,General,1,Options,Back

1,all,Activate,1,OK,Back

0,done

6,all,Personalise,1,Select,Back

6,one,Incoming,call alert, ,Caller groups,1,1,0,Select,Back

1,all,Ringing,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

1,all,Ring once,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Caller groups,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

36,one,Ringing,tone, ,Bee,1,1,0,Select,Back

1,all,Ring ring,1,OK,Back

0,done

1,all,Low,1,OK,Back
0,done
1,all,Intro,1,OK,Back
0,done
1,all,Bee,1,OK,Back
0,done
1,all,Fly,1,OK,Back
0,done
1,all,Mosquito,1,OK,Back
0,done
1,all,Do-mi-sol,1,OK,Back
0,done
1,all,Chase,1,OK,Back
0,done
1,all,Hunt,1,OK,Back
0,done
1,all,Grande valse,1,OK,Back
0,done
1,all,Etude,1,OK,Back
0,done
1,all,City bird,1,OK,Back
0,done
1,all,Going up,1,OK,Back
0,done
1,all,Polite,1,OK,Back
0,done
1,all,Attraction,1,OK,Back
0,done
1,all,Persuasion,1,OK,Back
0,done
1,all,Robo N1X,1,OK,Back
0,done
1,all,Scifi,1,OK,Back
0,done
1,all,That's it,1,OK,Back
0,done
1,all,Playground,1,OK,Back
0,done
1,all,Fuga,1,OK,Back
0,done
1,all,Badinerie,1,OK,Back
0,done
1,all,Menuet,1,OK,Back
0,done
1,all,Elise,1,OK,Back
0,done
1,all,Ode to Joy,1,OK,Back
0,done
1,all,Mozart 40,1,OK,Back
0,done
1,all,Piano Concerto,1,OK,Back
0,done
1,all,William Tell,1,OK,Back
0,done
1,all,Charleston,1,OK,Back
0,done
1,all,Samba,1,OK,Back
0,done
1,all,Polka,1,OK,Back
0,done
1,all,Helan,1,OK,Back
0,done
1,all,Orient,1,OK,Back
0,done
1,all,Tick tick,1,OK,Back
0,done

1,all,Jumping,1,OK,Back

0,done

1,all,Personal,1,OK,Back

0,done

5,one,Ringing,volume, ,Level 2,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Level 4,1,OK,Back

0,done

1,all,Level 5,1,OK,Back

0,done

5,one,Message,alert tone, ,Standard,1,1,0,Select,Back

1,all,No tone,1,OK,Back

0,done

1,all,Standard,1,OK,Back

0,done

1,all,Special,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

4,one,Keypad,tones, ,Level 1,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,one,Warning and,game tones, ,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,all,Silent,1,Options,Back

1,all,Activate,1,OK,Back

0,done

6,all,Personalise,1,Select,Back

6,one,Incoming,call alert, ,Caller groups,1,1,0,Select,Back

1,all,Ringing,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

1,all,Ring once,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Caller groups,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

36,one,Ringing,tone, ,Bee,1,1,0,Select,Back

1,all,Ring ring,1,OK,Back

0,done

1,all,Low,1,OK,Back

0,done

1,all,Intro,1,OK,Back

0,done

1,all,Bee,1,OK,Back

0,done

1,all,Fly,1,OK,Back

0,done

1,all,Mosquito,1,OK,Back

0,done

1,all,Do-mi-sol,1,OK,Back

0,done

1,all,Chase,1,OK,Back

0,done

1,all,Hunt,1,OK,Back

0,done

1,all,Grande valse,1,OK,Back

0,done

1,all,Etude,1,OK,Back

0,done

1,all,City bird,1,OK,Back

0,done

1,all,Going up,1,OK,Back

0,done

1,all,Polite,1,OK,Back

0,done

1,all,Attraction,1,OK,Back

0,done

1,all,Persuasion,1,OK,Back

0,done

1,all,Robo N1X,1,OK,Back

0,done

1,all,Scifi,1,OK,Back

0,done

1,all,That's it,1,OK,Back

0,done

1,all,Playground,1,OK,Back

0,done

1,all,Fuga,1,OK,Back

0,done

1,all,Badinerie,1,OK,Back

0,done

1,all,Menuet,1,OK,Back

0,done

1,all,Elise,1,OK,Back

0,done

1,all,Ode to Joy,1,OK,Back

0,done

1,all,Mozart 40,1,OK,Back

0,done

1,all,Piano Concerto,1,OK,Back

0,done

1,all,William Tell,1,OK,Back

0,done

1,all,Charleston,1,OK,Back

0,done

1,all,Samba,1,OK,Back

0,done

1,all,Polka,1,OK,Back

0,done

1,all,Helan,1,OK,Back

0,done

1,all,Orient,1,OK,Back

0,done

1,all,Tick tick,1,OK,Back

0,done

1,all,Jumping,1,OK,Back

0,done

1,all,Personal,1,OK,Back

0,done

5,one,Ringing,volume, ,Level

2,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Level 4,1,OK,Back

0,done

1,all,Level 5,1,OK,Back

0,done

5,one,Message,alert tone,

,Standard,1,1,0,Select,Back

1,all,No tone,1,OK,Back

0,done

1,all,Standard,1,OK,Back

0,done

1,all,Special,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

4,one,Keypad,tones, ,Level 1,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,one,Warning and,game tones,

,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,all,Meeting,1,Options,Back

1,all,Activate,1,OK,Back

0,done

6,all,Personalise,1,Select,Back

6,one,Incoming,call alert, ,Caller

groups,1,1,0,Select,Back

1,all,Ringing,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

1,all,Ring once,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Caller groups,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

36,one,Ringing,tone, ,Bee,1,1,0,Select,Back

1,all,Ring ring,1,OK,Back

0,done

1,all,Low,1,OK,Back

0,done

1,all,Intro,1,OK,Back

0,done

1,all,Bee,1,OK,Back

0,done

1,all,Fly,1,OK,Back

0,done

1,all,Mosquito,1,OK,Back

0,done

1,all,Do-mi-sol,1,OK,Back

0,done

1,all,Chase,1,OK,Back

0,done

1,all,Hunt,1,OK,Back

0,done

1,all,Grande valse,1,OK,Back

0,done

1,all,Etude,1,OK,Back

0,done

1,all,City bird,1,OK,Back

0,done

1,all,Going up,1,OK,Back

0,done

1,all,Polite,1,OK,Back

0,done

1,all,Attraction,1,OK,Back

0,done

1,all,Persuasion,1,OK,Back

0,done

1,all,Robo N1X,1,OK,Back

0,done

1,all,Scifi,1,OK,Back

0,done

1,all,That's it,1,OK,Back

0,done

1,all,Playground,1,OK,Back

0,done

1,all,Fuga,1,OK,Back

0,done

1,all,Badinerie,1,OK,Back

0,done

1,all,Menuet,1,OK,Back

0,done

1,all,Elise,1,OK,Back

0,done

1,all,Ode to Joy,1,OK,Back

0,done

1,all,Mozart 40,1,OK,Back

0,done

1,all,Piano Concerto,1,OK,Back

0,done

1,all,William Tell,1,OK,Back

0,done

1,all,Charleston,1,OK,Back

0,done

1,all,Samba,1,OK,Back

0,done

1,all,Polka,1,OK,Back

0,done

1,all,Helan,1,OK,Back

0,done

1,all,Orient,1,OK,Back

0,done

1,all,Tick tick,1,OK,Back

0,done

1,all,Jumping,1,OK,Back

0,done

1,all,Personal,1,OK,Back

0,done

5,one,Ringing,volume, ,Level
2,1,1,0,Select,Back
1,all,Level 1,1,OK,Back
0,done
1,all,Level 2,1,OK,Back
0,done
1,all,Level 3,1,OK,Back
0,done
1,all,Level 4,1,OK,Back
0,done
1,all,Level 5,1,OK,Back
0,done
5,one,Message,alert tone,
,Standard,1,1,0,Select,Back
1,all,No tone,1,OK,Back
0,done
1,all,Standard,1,OK,Back
0,done
1,all,Special,1,OK,Back
0,done
1,all,Beep once,1,OK,Back
0,done
1,all,Ascending,1,OK,Back
0,done
4,one,Keypad,tones, ,Level 1,1,1,0,Select,Back
1,all,Level 1,1,OK,Back
0,done
1,all,Level 2,1,OK,Back
0,done
1,all,Level 3,1,OK,Back
0,done
1,all,Off,1,OK,Back
0,done
2,one,Warning and,game tones,
,On,1,1,0,Select,Back
1,all,On,1,OK,Back
0,done
1,all,Off,1,OK,Back
0,done
2,all,Outdoor,1,Options,Back
1,all,Activate,1,OK,Back
0,done
6,all,Personalise,1,Select,Back
6,one,Incoming,call alert, ,Caller
groups,1,1,0,Select,Back
1,all,Ringing,1,OK,Back
0,done
1,all,Ascending,1,OK,Back
0,done
1,all,Ring once,1,OK,Back
0,done
1,all,Beep once,1,OK,Back
0,done
1,all,Caller groups,1,OK,Back
0,done
1,all,Off,1,OK,Back
0,done
36,one,Ringing,tone, ,Bee,1,1,0,Select,Back
1,all,Ring ring,1,OK,Back
0,done
1,all,Low,1,OK,Back
0,done
1,all,Intro,1,OK,Back
0,done
1,all,Bee,1,OK,Back
0,done

1,all,Fly,1,OK,Back
0,done
1,all,Mosquito,1,OK,Back
0,done
1,all,Do-mi-sol,1,OK,Back
0,done
1,all,Chase,1,OK,Back
0,done
1,all,Hunt,1,OK,Back
0,done
1,all,Grande valse,1,OK,Back
0,done
1,all,Etude,1,OK,Back
0,done
1,all,City bird,1,OK,Back
0,done
1,all,Going up,1,OK,Back
0,done
1,all,Polite,1,OK,Back
0,done
1,all,Attraction,1,OK,Back
0,done
1,all,Persuasion,1,OK,Back
0,done
1,all,Robo N1X,1,OK,Back
0,done
1,all,Scifi,1,OK,Back
0,done
1,all,That's it,1,OK,Back
0,done
1,all,Playground,1,OK,Back
0,done
1,all,Fuga,1,OK,Back
0,done
1,all,Badinerie,1,OK,Back
0,done
1,all,Menuet,1,OK,Back
0,done
1,all,Elise,1,OK,Back
0,done
1,all,Ode to Joy,1,OK,Back
0,done
1,all,Mozart 40,1,OK,Back
0,done
1,all,Piano Concerto,1,OK,Back
0,done
1,all,William Tell,1,OK,Back
0,done
1,all,Charleston,1,OK,Back
0,done
1,all,Samba,1,OK,Back
0,done
1,all,Polka,1,OK,Back
0,done
1,all,Helan,1,OK,Back
0,done
1,all,Orient,1,OK,Back
0,done
1,all,Tick tick,1,OK,Back
0,done
1,all,Jumping,1,OK,Back
0,done
1,all,Personal,1,OK,Back
0,done
5,one,Ringing,volume, ,Level
2,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Level 4,1,OK,Back

0,done

1,all,Level 5,1,OK,Back

0,done

5,one,Message,alert tone,

,Standard,1,1,0,Select,Back

1,all,No tone,1,OK,Back

0,done

1,all,Standard,1,OK,Back

0,done

1,all,Special,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

4,one,Keypad,tones, ,Level 1,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,one,Warning and,game tones,

,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,all,Pager,1,Options,Back

1,all,Activate,1,OK,Back

0,done

6,all,Personalise,1,Select,Back

6,one,Incoming,call alert, ,Caller

groups,1,1,0,Select,Back

1,all,Ringing,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

1,all,Ring once,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Caller groups,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

36,one,Ringing,tone, ,Bee,1,1,0,Select,Back

1,all,Ring ring,1,OK,Back

0,done

1,all,Low,1,OK,Back

0,done

1,all,Intro,1,OK,Back

0,done

1,all,Bee,1,OK,Back

0,done

1,all,Fly,1,OK,Back

0,done

1,all,Mosquito,1,OK,Back

0,done

1,all,Do-mi-sol,1,OK,Back

0,done

1,all,Chase,1,OK,Back

0,done

1,all,Hunt,1,OK,Back

0,done

1,all,Grande valse,1,OK,Back

0,done

1,all,Etude,1,OK,Back

0,done

1,all,City bird,1,OK,Back

0,done

1,all,Going up,1,OK,Back

0,done

1,all,Polite,1,OK,Back

0,done

1,all,Attraction,1,OK,Back

0,done

1,all,Persuasion,1,OK,Back

0,done

1,all,Robo N1X,1,OK,Back

0,done

1,all,Scifi,1,OK,Back

0,done

1,all,That's it,1,OK,Back

0,done

1,all,Playground,1,OK,Back

0,done

1,all,Fuga,1,OK,Back

0,done

1,all,Badinerie,1,OK,Back

0,done

1,all,Menuet,1,OK,Back

0,done

1,all,Elise,1,OK,Back

0,done

1,all,Ode to Joy,1,OK,Back

0,done

1,all,Mozart 40,1,OK,Back

0,done

1,all,Piano Concerto,1,OK,Back

0,done

1,all,William Tell,1,OK,Back

0,done

1,all,Charleston,1,OK,Back

0,done

1,all,Samba,1,OK,Back

0,done

1,all,Polka,1,OK,Back

0,done

1,all,Helan,1,OK,Back

0,done

1,all,Orient,1,OK,Back

0,done

1,all,Tick tick,1,OK,Back

0,done

1,all,Jumping,1,OK,Back

0,done

1,all,Personal,1,OK,Back

0,done

5,one,Ringing,volume, ,Level

2,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Level 4,1,OK,Back

0,done

1,all,Level 5,1,OK,Back

0,done

5,one,Message,alert tone,

,Standard,1,1,0,Select,Back

1,all,No tone,1,OK,Back

0,done

1,all,Standard,1,OK,Back

0,done

1,all,Special,1,OK,Back

0,done

1,all,Beep once,1,OK,Back

0,done

1,all,Ascending,1,OK,Back

0,done

4,one,Keypad,tones, ,Level 1,1,1,0,Select,Back

1,all,Level 1,1,OK,Back

0,done

1,all,Level 2,1,OK,Back

0,done

1,all,Level 3,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,one,Warning and,game tones,

,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

7,one,Settings, , , ,0,1,1,Select,Exit

2,one,Alarm,clock, ,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

0,one,Auto update of,date and time, ,

,0,1,0,Select,Back

0,one,Clock, , ,02:29pm,1,1,0,Select,Back

5,one,Call,settings, , ,0,1,0,Select,Back

2,one,Anykey,answer, ,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,one,Automatic,redial, ,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

2,one,Speed,dialling, ,On,1,1,0,Select,Back

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

3,one,Call,waiting, , ,1,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Own number,sending, ,
,1,1,0,Select,Back
1,all,Preset,1,OK,Back
0,done
1,all,On,1,OK,Back
0,done
1,all,Off,1,OK,Back
0,done
6,one,Phone,settings, , ,0,1,0,Select,Back
9,one,Language, , ,English,1,1,0,Select,Back
0,all,Automatic,1,OK,Back
0,all,English,1,OK,Back
0,all,Deutsch,1,OK,Back
0,all,Francais,1,OK,Back
0,all,Nederlands,1,OK,Back
0,all,Italiano,1,OK,Back
0,all,Dansk,1,OK,Back
0,all,Svenska,1,OK,Back
0,all,Norsk,1,OK,Back
2,one,Cell info,display, ,Off,1,1,0,Select,Back
1,all,On,1,OK,Back
0,done
1,all,Off,1,OK,Back
0,done
0,one,Welcome,note, , ,0,1,0,Select,Back
0,one,Confirm,SIM service,actions,
,0,1,0,Select,Back
0,one,List of own,numbers, ,
,0,1,0,Select,Back
2,one,Network,connection,
,Automatic,1,1,0,Select,Back
1,all,Automatic,1,OK,Back
0,done
1,all,Manual,1,OK,Back
0,done
6,one,Security,settings, , ,0,1,0,Select,Back
0,one,PIN code,request, ,On,0,1,0,Select,Back
6,one,Call barring,service, , ,0,1,0,Select,Back
3,one,Outgoing,calls, , ,0,1,0,Select,Back
1,all,Activate,1,OK,Back
0,done
1,all,Cancel,1,OK,Back
0,done
0,all,Status,1,OK,Back
3,one,International,calls, , ,0,1,0,Select,Back
1,all,Activate,1,OK,Back
0,done
1,all,Cancel,1,OK,Back
0,done
0,all,Status,1,OK,Back
3,one,International,but home country, ,
,0,1,0,Select,Back
1,all,Activate,1,OK,Back
0,done
1,all,Cancel,1,OK,Back
0,done
0,all,Status,1,OK,Back
3,one,Incoming,calls, , ,0,1,0,Select,Back
1,all,Activate,1,OK,Back
0,done
1,all,Cancel,1,OK,Back
0,done
0,all,Status,1,OK,Back
3,one,Incoming calls,if abroad, ,
,0,1,0,Select,Back
1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

1,one,Cancel all,barrings, , ,0,1,0,Select,Back

0,done

0,one,Fixed,dialling, ,Off,0,1,0,Select,Back

3,one,Closed,user group,

,Preset,1,1,0,Select,Back

1,all,Preset,1,OK,Back

0,done

1,all,On,1,OK,Back

0,done

1,all,Off,1,OK,Back

0,done

0,one,Security,level, ,None,0,1,0,Select,Back

4,one,Change,access codes, ,

,1,1,0,Select,Back

0,one,Change,security code, ,

,1,1,0,Select,Back

0,one,Change,PIN code, , ,1,1,0,Select,Back

0,one,Change,PIN2 code, , ,1,1,0,Select,Back

0,one,Change,barring password, ,

,1,1,0,Select,Back

0,one,Restore,factory settings, ,

,0,1,0,Select,Back

8,one,Call divert, , , ,0,1,1,Select,Exit

3,one,Divert voice calls,without ringing, ,

,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Divert,when busy, , ,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Divert when,not answered, ,

,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Divert when,phone off 1, ,

,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Divert when,phone off 2, ,

,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Divert all,fax calls, , ,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Divert all,data calls, , ,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

3,one,Cancel all,diverts, , ,0,1,0,Select,Back

1,all,Activate,1,OK,Back

0,done

1,all,Cancel,1,OK,Back

0,done

0,all,Status,1,OK,Back

4,one,Games, , , ,0,1,1,Select,Exit

0,one,Memory, , , ,1,1,0,Select,Back

0,one,Snake, , , ,1,1,0,Select,Back

0,one,Logic, , , ,1,1,0,Select,Back

0,one,Rotation, , , ,1,1,0,Select,Back

0,one,Calculator, , , ,0,1,1,Select,Exit

0,one,Calendar, , , ,0,1,1,Select,Exit

0,one,Infrared, , , ,0,1,1,Select,Exit

2,one,Services, , , ,0,1,1,Select,Exit

0,one,Personal,bookmarks, ,

,1,1,0,Select,Back

0,one,Operator,services, , ,1,1,0,Select,Back

0,one,SIM services, , , ,0,1,1,Select,Exit

Appendix C

Experiment 2 Data

C.1 Instructions

1. The aim of this experiment is to investigate the use of non-speech audio to support navigation in hierarchical menus with limited graphical display. In particular, we would like to find out if sound can be used to help novice users learn the structure of a menu.

2. Your task will be to perform a sequence of actions with the simulation of the mobile phone on the left of this window.

3. Each task description will show in this window. When you have completed a task, press the space bar to read the instructions for the next task. You don't need to click on the current window to press the space bar. You can always leave the phone window active. If you accidentally press the space bar twice, please let me know.

4. A description of some of the phone features will be provided to you as part of the task description when necessary. However, some of the tasks may require you to undertake an exhaustive search into the menu. It might take you a while to find some of the phone functions.

5. The red button of the phone allows you to jump to the top level menu from anywhere in the menu. For some of the tasks, you will be asked to use this button. In the rest of the experiment you are free to use it or not to complete a task.

6. There are essentially 3 types of tasks that you will be asked to complete... A) Tasks in which you have to ACCOMPLISH an an action implemented in the menu. In this case,

press the space bar after you think you have completed the setup successfully. B) Tasks in which you have to FIND an option or function in the menu. When you think that you have reached the required option or function, you can carry on to the next task by pressing the space bar. C) Tasks for which you have to answer a question. In that case, just write down the answer on the sheet of paper provided.

7. At the end of the experiment, you will be asked to fill a 'workload' form, to evaluate the mental, physical efforts, annoyance, etc, that you have expanded during the experiment. Please read NOW the description of this workload form.

8. Thank you in advance for your participation. Don't hesitate to ask me if anything is not quite clear at any point of the experiment. You may now press the space bar when you are ready to start the experiment ...

C.2 List of Tasks

1. Scan all main menus twice.

2. In the 'Profiles' menu, scan all the submenus until you come back to the first one twice. Just visit the 'first level' submenus, you don't have to go deeper in the menu.

3. Press the red key to go to the top level menu.

4. In the 'Settings' menu, scan all the submenus until you come back to the first one twice. Just visit the 'first level' submenus, you don't have to go deeper in the menu.

5. Press the red key to go to the top level menu.

6. In the 'Call divert' menu, scan all the submenus twice until you come back to the first one. Just visit the 'first level' submenus, you don't have to go deeper in the menu.

7. Press the red key to go to the top level menu.

8. Activate the outdoor profile.

9. In the meeting profile, set the ringing tone to 'Samba'.

10. Press the red key to go to the top level menu

11. Change the ringing volume to 'level 3' in the same profile (meeting).

12. In the general profile, set the ringing tone to 'Fly'.
13. Press the red key to go back to the top level menu.
14. In the same profile (general), set the ringing volume to 'level 1'.
15. Press the red key to go back to the top level menu.
16. Still in the same profile (general), set the keypad tone volume to 'level 3'.
17. Press the red key to go back to the top level menu.
18. This phone allows you to divert unwanted calls. Set it up so that all calls are diverted when busy.
19. Press the red key to go back to the top level menu.
20. Set it up to divert all fax calls.
21. When you switch on the phone, a brief welcome note shows on the screen. Try and find where in the menu this welcome note can be configured.
22. The time is shown on the top level menu, as you can see by pressing the red button. Find where the clock can be configured. Note: In this simulation the time cannot be configured; when you reach the 'clock' function, just proceed to the next task.
23. This phone keeps a record of the calls made, received, etc, until they are erased. Check that there is one dialled number recorded and then, erase it.
24. Change the ringing tone of the meeting profile to 'bee'.
25. When the phone is ringing, you can take the call by pressing a particular key, or any key at all. This is configurable. Configure it so that you can NOT answer a call by pressing any key.
26. When a call is given from this phone, its own phone number may or may not be sent. Switch off the 'Own number sending' function.
27. Find where the language can be configured.
28. The phone features a call barring service. Activate call barrings for international calls.
29. Press the red key to go back to the top level menu.

30. Activate call barrings for incoming calls.

31. Switch on the 'Speed dialling' function.

32. Earlier in the experiment, you have setup some 'call diverts'. Now Cancel them all.

33. In the 'general' profile, switch off the warning and games tones.

34. Erase the call received from Steve.

35. What is the duration of the last call.

36. Clear the timers, e.g. the record of all call durations.

37. Divert all data calls.

38. Divert calls when busy.

39. Find the clock again e.g., the place where the time can be configured.

40. What is the volume of the keypad tones in the silent profile?

41. Retrieve the place in the menu where the language can be configured.

42. How many calls are listed in the 'received calls' menu.

43. Cancel all call barrings.

44. Cancel all call diverts.

45. Is the Game menu available e.g., does it have any submenu?

46. Is the calculator available?

47. Is the calendar available?

48. Find the clock again (the place where the time could be configured).

49. Is there a 'Template' submenu in the 'Messages' menu?

50. What is the ringing tone in the silent profile?

51. Activate the call barring service for all incoming calls.

52. Find the welcome note again.

53. Set speed dialling off.

54. Clear timers.
55. Cancel all call barrings.
56. Cancel all call diverts.

C.3 Questionnaire answers

These are the answers given by the eleven respondants from Group 1 to the following questions:

C.3.1 Questions

- Do you think that your overall performance benefitted from the sounds?
    1. Parts of it when remembering number of different profiles.
    2. Yes, although more towards the end.
    3. Well, it would have been worse without it.
    4. Slightly but more so for the ???? test.
    5. Yes.
    6. The sounds were beneficial in that they provided additional feedback as a result of interaction with the keyboard. However I spent more time interpreting visual clues (slotted bar on RHS), which is probably typical of computer scientists.
    7. Not significantly.
    8. Not sure.
    9. Yes.
    10. Hard to say.
    11. Yes.
- Where the sounds meaningful to you?
    1. Yes, usually when the pitch of the sound changed dramatically it was obvious I was back to the first option in the menu.
    2. Yes, they provided a rhythm to selecting a menu.
    3. I think you can tell quicker that you're looking at something you already looked at.

4. Only in terms of pitch.
5. Yes, in terms of identifying your position in the hierarchy.
6. The sounds were pleasant, but did not convey any structured information to me. I would probably only recognise the "DONE" audio clue.
7. Yes.
8. Some.
9. Yes.
10. Yes.
11. Yes. Reduced the need to keep track of the menu items that I had scanned. Helped locate headings in the main menu.

- What kind of information was provided by the sounds?

1. When you were at the first option in the menu. When I clicked the red button to go back to the top.
2. A kind of timing to select the menus. It also meant I was unlikely to go too fast and make a mistake.
3. see above.
4. Position in the current menu/submenu.
5. Finding my place in the structure and whether there were submenus under my current position.
6. Non that consciously registered in my mind.
7. Your current position in the menu. If a task was completed successfully. If a menu has a submenu.
8. Type of menu. Option number in a menu.
9. Whether I had looped around an entire menu. Remembered how to find certain menu items.
10. Position in menu.
11. Priority. Numerical.

- What did you dislike most about the sounds?

1. Some of them started to annoy me.
2. The majority were soft and some of hem could have been in a different style.
3. Nothing.
4. They don't relate to the menu option and I didn't practice enough to learn them to the menu options.
5. Nothing.
6. I found the sounds rather pleasant (unobstrusive) and soothing.
7. Some were rather intrusive.
8. Disturbing.
9. Nothing.
10. Perhaps the sounds were too complex to learn easily but ...
11. Some provided no info (e.g. in change tone). Not immediatly apparent why sounds differ in main menu. No consistency in their use.

- What did you like about them?

1. Some sounds, like the one you hear when you click the red button, helped me determine if I had pressed it or missed it.
2. They were very pleasant and when moving in a menu provided a pleasant melody.
3. See question 2.
4. Short, not too loud, and not too irritating (as a simulated voice might be).
5. Some wre funny!! Useful clues for the level of the hierarchy.
6. Mellow character.
7. It is good to know were you are in the menu - the decreasing pitch was good. It's also good to know when a task was done successfully.
8. While scanning options in a menu, you can hear when you are back to the beginning.
9. Gave me some more feedback as I found the directory structure quite hard to navigate and remember where to find items.
10. They were musical - did not sound computer generated.
11. See question2. Made interface more lively (?), more sophisitcated.

- Any other comments?...

  1. None.
  2. I think it has been more fun than it would have been using the menus without the osunds.
  3. I was rubbish at it but the sounds did help.
  4. Thank goodness I don't have a mobile phone.
  5. The blind task was cruel!!
  6. One or two instructions could have been worded more clearly, e.g. "until you come back to the first menu twice", which was later expressed differently and unambiguously.
  7. The sounds were rather annoying tostart with, but you get (a little bit) used to them after a while.
  8. I do not like mobile phone!
  9. There was a lot on the menu and overall I just found coming to terms with all the items difficult.
  10. None.
  11. Some tone increment would be usefule in all menus, except for the very large ones (e.g. main menu).

C.4 Raw Data

tasks/subjects	**s1**	**s2**	**s3**	**s4**	**s5**	**s6**	**s7**	**s8**	**s9**	**s10**	**s11**	**s12**	**sum**	**max**
8	13	11	11	20	11	11	11	18	11	11	11	11	150	20
9	37	41	41	44	35	76	38	107	37	39	206	37	738	206
11	16	16	16	16	25	16	16	35	16	16	16	16	220	35
12	17	19	15	17	42	16	13	24	14	13	13	13	216	42
14	14	14	16	14	14	14	14	37	14	14	18	19	202	37
16	19	17	19	17	21	17	17	26	17	17	19	19	225	26
18	18	15	11	11	11	67	18	15	18	15	22	11	232	67
20	14	17	14	16	14	14	14	17	14	14	20	14	182	20
21	73	26	26	131	24	27	77	19	25	236	37	27	728	236
22	17	7	8	8	7	7	4	17	7	7	22	7	118	22
23	40	20	13	116	14	16	28	25	23	97	31	14	437	116
24	41	32	27	27	16	41	27	36	27	35	31	27	367	41
25	15	46	16	24	20	92	14	24	16	20	26	29	342	92
26	45	35	29	22	21	10	24	20	21	10	72	12	321	72
27	4	20	25	17	11	106	51	53	11	76	14	19	407	106
28	37	39	59	91	26	43	304	44	49	41	97	69	899	304
30	50	19	21	19	19	25	35	65	24	24	21	19	341	65
31	11	0	32	28	27	19	39	28	17	20	26	24	271	39
32	31	23	17	70	97	11	32	18	69	95	31	12	506	97
33	19	25	20	18	18	18	18	19	19	19	22	17	232	25
34	49	68	17	15	14	21	29	62	16	13	15	48	367	68
35	23	9	27	59	12	8	34	49	10	9	5	23	268	59
36	8	16	9	8	11	21	8	23	5	7	15	8	139	23
37	56	20	15	16	22	16	21	17	20	22	22	15	262	56
38	8	8	7	7	8	7	7	8	10	8	8	7	93	10
39	6	8	7	7	4	7	9	14	9	7	7	9	94	14
40	9	17	22	15	13	13	13	28	17	13	15	13	188	28
41	11	20	11	10	26	12	80	14	14	33	21	23	275	80
42	16	10	8	8	8	14	15	20	10	47	11	8	175	47
43	56	87	54	21	67	43	0	39	32	31	29	60	519	87
44	8	12	13	16	11	16	50	13	45	36	22	13	255	50
45	8	7	1	6	1	8	4	1	8	8	6	1	59	8
46	1	1	1	1	1	7	5	8	7	1	1	1	35	8
47	1	1	1	1	1	1	8	7	1	0	1	1	24	8
48	7	17	6	6	7	7	7	14	6	5	9	6	97	17
50	9	13	11	9	10	10	10	15	12	9	10	9	127	15
51	30	60	69	20	26	20	0	30	25	32	35	38	385	69
52	8	15	6	22	12	12	21	14	9	18	14	9	160	22
53	17	25	19	10	18	18	20	19	11	47	23	12	239	47
54	47	147	307	43	11	16	115	34	196	123	118	23	1180	307
55	41	22	39	19	14	25	0	32	19	31	87	47	376	87
56	9	30	13	11	12	16	23	17	10	23	12	13	189	30
sum	959	1055	1099	1056	782	964	1273	1125	941	1342	1241	803	12640	

Number of keystrokes taken by Group 1 to complete the non-trial tasks.

tasks/subjects	s13	s14	s15	s16	s17	s18	s19	20	s21	s22	s23	s24	sum	max
8	11	11	11	11	56	16	11	20	17	21	11	11	207	56
9	63	51	39	128	138	71	47	37	124	88	74	37	897	138
11	21	16	18	28	72	38	16	16	16	27	17	18	303	72
12	29	13	15	15	32	13	19	13	14	18	15	15	211	32
14	14	14	22	24	18	27	14	14	14	29	14	16	220	29
16	15	17	19	22	19	24	17	19	25	32	17	23	249	32
18	18	11	11	13	51	14	18	11	37	22	65	11	282	65
20	14	15	14	32	14	19	14	20	14	16	14	14	200	32
21	187	26	32	160	115	82	21	36	14	31	15	15	734	187
22	10	7	0	4	6	3	7	28	16	8	7	20	116	28
23	21	42	26	67	16	13	22	21	13	34	19	15	309	67
24	26	43	29	60	20	19	31	35	47	57	47	32	446	60
25	19	14	43	25	23	72	31	45	23	66	19	41	421	72
26	48	65	10	12	7	12	12	54	19	49	30	12	330	65
27	157	21	24	10	17	27	8	22	52	212	11	29	590	212
28	250	29	70	126	33	136	62	21	143	68	194	33	1165	250
30	0	21	37	19	41	25	19	26	35	58	80	31	392	80
31	66	17	37	18	22	49	12	13	31	45	23	26	359	66
32	23	16	110	57	22	76	56	39	73	212	17	13	714	212
33	18	20	23	17	16	28	23	19	22	22	20	40	268	40
34	16	33	67	127	77	23	25	53	17	112	16	29	595	127
35	45	9	5	12	7	7	53	8	6	30	5	5	192	53
36	41	22	8	4	20	5	10	5	8	17	16	5	161	41
37	18	12	27	16	6	46	16	18	16	13	16	15	219	46
38	8	10	10	9	8	8	9	0	8	10	8	8	96	10
39	11	8	41	8	7	4	8	8	14	12	7	4	132	41
40	33	13	25	38	12	38	16	12	21	33	45	11	297	45
41	16	10	146	157	10	9	24	39	16	126	11	9	573	157
42	48	20	46	6	8	18	16	19	10	70	24	15	300	70
43	0	16	70	27	49	27	37	0	41	15	58	54	394	70
44	31	8	87	48	16	48	8	80	60	16	56	15	473	87
45	2	3	0	13	3	17	6	8	51	7	9	6	125	51
46	5	1	2	8	1	8	1	1	3	8	1	3	42	8
47	6	1	1	9	1	14	1	1	1	3	1	5	44	14
48	14	7	6	6	7	6	16	12	6	0	7	16	103	16
50	34	14	0	19	10	16	17	0	13	16	0	9	148	34
51	0	21	36	36	20	29	30	91	24	0	97	21	405	97
52	36	23	19	15	12	33	9	13	6	0	10	18	194	36
53	29	18	33	15	19	41	15	48	34	34	15	19	320	48
54	38	52	92	158	71	69	32	29	196	38	97	18	890	196
55	0	18	20	50	22	24	16	54	41	35	0	22	302	54
56	16	8	17	3	13	25	12	0	35	34	0	16	179	35
sum	1457	796	1348	1632	1137	1279	837	1008	1376	1744	1208	775	14597	

Number of keystrokes taken by Group 2 to complete the non-trial tasks.

tasks/subjects	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	sum	max
8	0.23	0.20	0.20	0.36	0.20	0.20	0.20	0.32	0.20	0.20	0.20	0.20	2.68	0.36
9	0.18	0.20	0.20	0.21	0.17	0.37	0.18	0.52	0.18	0.19	1.00	0.18	3.58	1.00
11	0.22	0.22	0.22	0.22	0.35	0.22	0.22	0.49	0.22	0.22	0.22	0.22	3.06	0.49
12	0.40	0.45	0.36	0.40	1.00	0.38	0.31	0.57	0.33	0.31	0.31	0.31	5.14	1.00
14	0.38	0.38	0.43	0.38	0.38	0.38	0.38	1.00	0.38	0.38	0.49	0.51	5.46	1.00
16	0.59	0.53	0.59	0.53	0.66	0.53	0.53	0.81	0.53	0.53	0.59	0.59	7.03	0.81
18	0.27	0.22	0.16	0.16	0.16	1.00	0.27	0.22	0.27	0.22	0.33	0.16	3.46	1.00
20	0.44	0.53	0.44	0.50	0.44	0.44	0.44	0.53	0.44	0.44	0.63	0.44	5.69	0.63
21	0.31	0.11	0.11	0.56	0.10	0.11	0.33	0.08	0.11	1.00	0.16	0.11	3.08	1.00
22	0.61	0.25	0.29	0.29	0.25	0.25	0.14	0.61	0.25	0.25	0.79	0.25	4.21	0.79
23	0.34	0.17	0.11	1.00	0.12	0.14	0.24	0.22	0.20	0.84	0.27	0.12	3.77	1.00
24	0.68	0.53	0.45	0.45	0.27	0.68	0.45	0.60	0.45	0.58	0.52	0.45	6.12	0.68
25	0.16	0.50	0.17	0.26	0.22	1.00	0.15	0.26	0.17	0.22	0.28	0.32	3.72	1.00
26	0.63	0.49	0.40	0.31	0.29	0.14	0.33	0.28	0.29	0.14	1.00	0.17	4.46	1.00
27	0.02	0.09	0.12	0.08	0.05	0.50	0.24	0.25	0.05	0.36	0.07	0.09	1.92	0.50
28	0.12	0.13	0.19	0.30	0.09	0.14	1.00	0.14	0.16	0.13	0.32	0.23	2.96	1.00
30	0.63	0.24	0.26	0.24	0.24	0.31	0.44	0.81	0.30	0.30	0.26	0.24	4.26	0.81
31	0.17	0.00	0.48	0.42	0.41	0.29	0.59	0.42	0.26	0.30	0.39	0.36	4.11	0.59
32	0.15	0.11	0.08	0.33	0.46	0.05	0.15	0.08	0.33	0.45	0.15	0.06	2.39	0.46
33	0.48	0.63	0.50	0.45	0.45	0.45	0.45	0.48	0.48	0.48	0.55	0.43	5.80	0.63
34	0.39	0.54	0.13	0.12	0.11	0.17	0.23	0.49	0.13	0.10	0.12	0.38	2.89	0.54
35	0.39	0.15	0.46	1.00	0.20	0.14	0.58	0.83	0.17	0.15	0.08	0.39	4.54	1.00
36	0.20	0.39	0.22	0.20	0.27	0.51	0.20	0.56	0.12	0.17	0.37	0.20	3.39	0.56
37	1.00	0.36	0.27	0.29	0.39	0.29	0.38	0.30	0.36	0.39	0.39	0.27	4.68	1.00
38	0.80	0.80	0.70	0.70	0.80	0.70	0.70	0.80	1.00	0.80	0.80	0.70	9.30	1.00
39	0.15	0.20	0.17	0.17	0.10	0.17	0.22	0.34	0.22	0.17	0.17	0.22	2.29	0.34
40	0.20	0.38	0.49	0.33	0.29	0.29	0.29	0.62	0.38	0.29	0.33	0.29	4.18	0.62
41	0.07	0.13	0.07	0.06	0.17	0.08	0.51	0.09	0.09	0.21	0.13	0.15	1.75	0.51
42	0.23	0.14	0.11	0.11	0.11	0.20	0.21	0.29	0.14	0.67	0.16	0.11	2.50	0.67
43	0.64	1.00	0.62	0.24	0.77	0.49	0.00	0.45	0.37	0.36	0.33	0.69	5.97	1.00
44	0.09	0.14	0.15	0.18	0.13	0.18	0.57	0.15	0.52	0.41	0.25	0.15	2.93	0.57
45	0.16	0.14	0.02	0.12	0.02	0.16	0.08	0.02	0.16	0.16	0.12	0.02	1.16	0.16
46	0.13	0.13	0.13	0.13	0.13	0.88	0.63	1.00	0.88	0.13	0.13	0.13	4.38	1.00
47	0.07	0.07	0.07	0.07	0.07	0.07	0.57	0.50	0.07	0.00	0.07	0.07	1.71	0.57
48	0.41	1.00	0.35	0.35	0.41	0.41	0.41	0.82	0.35	0.29	0.53	0.35	5.71	1.00
50	0.26	0.38	0.32	0.26	0.29	0.29	0.29	0.44	0.35	0.26	0.29	0.26	3.74	0.44
51	0.31	0.62	0.71	0.21	0.27	0.21	0.00	0.31	0.26	0.33	0.36	0.39	3.97	0.71
52	0.22	0.42	0.17	0.61	0.33	0.33	0.58	0.39	0.25	0.50	0.39	0.25	4.44	0.61
53	0.35	0.52	0.40	0.21	0.38	0.38	0.42	0.40	0.23	0.98	0.48	0.25	4.98	0.98
54	0.15	0.48	1.00	0.14	0.04	0.05	0.37	0.11	0.64	0.40	0.38	0.07	3.84	1.00
55	0.47	0.25	0.45	0.22	0.16	0.29	0.00	0.37	0.22	0.36	1.00	0.54	4.32	1.00
56	0.26	0.86	0.37	0.31	0.34	0.46	0.66	0.49	0.29	0.66	0.34	0.37	5.40	0.86
sum	13.95	15.06	13.16	13.49	12.07	14.32	14.94	18.46	12.77	15.33	15.75	11.68	170.96	14.25

Normalised number of keystrokes taken by Group 1 to complete the non-trial tasks.

tasks/subjects	s13	s14	s15	s16	s17	s18	s19	20	s21	s22	s23	s24	sum	max
8	0.20	0.20	0.20	0.20	1.00	0.29	0.20	0.36	0.30	0.38	0.20	0.20	3.70	1.00
9	0.31	0.25	0.19	0.62	0.67	0.34	0.23	0.18	0.60	0.43	0.36	0.18	4.35	0.67
11	0.29	0.22	0.25	0.39	1.00	0.53	0.22	0.22	0.22	0.38	0.24	0.25	4.21	1.00
12	0.69	0.31	0.36	0.36	0.76	0.31	0.45	0.31	0.33	0.43	0.36	0.36	5.02	0.76
14	0.38	0.38	0.59	0.65	0.49	0.73	0.38	0.38	0.38	0.78	0.38	0.43	5.95	0.78
16	0.47	0.53	0.59	0.69	0.59	0.75	0.53	0.59	0.78	1.00	0.53	0.72	7.78	1.00
18	0.27	0.16	0.16	0.19	0.76	0.21	0.27	0.16	0.55	0.33	0.97	0.16	4.21	0.97
20	0.44	0.47	0.44	1.00	0.44	0.59	0.44	0.63	0.44	0.50	0.44	0.44	6.25	1.00
21	0.79	0.11	0.14	0.68	0.49	0.35	0.09	0.15	0.06	0.13	0.06	0.06	3.11	0.79
22	0.36	0.25	0.00	0.14	0.21	0.11	0.25	1.00	0.57	0.29	0.25	0.71	4.14	1.00
23	0.18	0.36	0.22	0.58	0.14	0.11	0.19	0.18	0.11	0.29	0.16	0.13	2.66	0.58
24	0.43	0.72	0.48	1.00	0.33	0.32	0.52	0.58	0.78	0.95	0.78	0.53	7.43	1.00
25	0.21	0.15	0.47	0.27	0.25	0.78	0.34	0.49	0.25	0.72	0.21	0.45	4.58	0.78
26	0.67	0.90	0.14	0.17	0.10	0.17	0.17	0.75	0.26	0.68	0.42	0.17	4.58	0.90
27	0.74	0.10	0.11	0.05	0.08	0.13	0.04	0.10	0.25	1.00	0.05	0.14	2.78	1.00
28	0.82	0.10	0.23	0.41	0.11	0.45	0.20	0.07	0.47	0.22	0.64	0.11	3.83	0.82
30	0.00	0.26	0.46	0.24	0.51	0.31	0.24	0.33	0.44	0.73	1.00	0.39	4.90	1.00
31	1.00	0.26	0.56	0.27	0.33	0.74	0.18	0.20	0.47	0.68	0.35	0.39	5.44	1.00
32	0.11	0.08	0.52	0.27	0.10	0.36	0.26	0.18	0.34	1.00	0.08	0.06	3.37	1.00
33	0.45	0.50	0.58	0.43	0.40	0.70	0.58	0.48	0.55	0.55	0.50	1.00	6.70	1.00
34	0.13	0.26	0.53	1.00	0.61	0.18	0.20	0.42	0.13	0.88	0.13	0.23	4.69	1.00
35	0.76	0.15	0.08	0.20	0.12	0.12	0.90	0.14	0.10	0.51	0.08	0.08	3.25	0.90
36	1.00	0.54	0.20	0.10	0.49	0.12	0.24	0.12	0.20	0.41	0.39	0.12	3.93	1.00
37	0.32	0.21	0.48	0.29	0.11	0.82	0.29	0.32	0.29	0.23	0.29	0.27	3.91	0.82
38	0.80	1.00	1.00	0.90	0.80	0.80	0.90	0.00	0.80	1.00	0.80	0.80	9.60	1.00
39	0.27	0.20	1.00	0.20	0.17	0.10	0.20	0.20	0.34	0.29	0.17	0.10	3.22	1.00
40	0.73	0.29	0.56	0.84	0.27	0.84	0.36	0.27	0.47	0.73	1.00	0.24	6.60	1.00
41	0.10	0.06	0.93	1.00	0.06	0.06	0.15	0.25	0.10	0.80	0.07	0.06	3.65	1.00
42	0.69	0.29	0.66	0.09	0.11	0.26	0.23	0.27	0.14	1.00	0.34	0.21	4.29	1.00
43	0.00	0.18	0.80	0.31	0.56	0.31	0.43	0.00	0.47	0.17	0.67	0.62	4.53	0.80
44	0.36	0.09	1.00	0.55	0.18	0.55	0.09	0.92	0.69	0.18	0.64	0.17	5.44	1.00
45	0.04	0.06	0.00	0.25	0.06	0.33	0.12	0.16	1.00	0.14	0.18	0.12	2.45	1.00
46	0.63	0.13	0.25	1.00	0.13	1.00	0.13	0.13	0.38	1.00	0.13	0.38	5.25	1.00
47	0.43	0.07	0.07	0.64	0.07	1.00	0.07	0.07	0.07	0.21	0.07	0.36	3.14	1.00
48	0.82	0.41	0.35	0.35	0.41	0.35	0.94	0.71	0.35	0.00	0.41	0.94	6.06	0.94
50	1.00	0.41	0.00	0.56	0.29	0.47	0.50	0.00	0.38	0.47	0.00	0.26	4.35	1.00
51	0.00	0.22	0.37	0.37	0.21	0.30	0.31	0.94	0.25	0.00	1.00	0.22	4.18	1.00
52	1.00	0.64	0.53	0.42	0.33	0.92	0.25	0.36	0.17	0.00	0.28	0.50	5.39	1.00
53	0.60	0.38	0.69	0.31	0.40	0.85	0.31	1.00	0.71	0.71	0.31	0.40	6.67	1.00
54	0.12	0.17	0.30	0.51	0.23	0.22	0.10	0.09	0.64	0.12	0.32	0.06	2.90	0.64
55	0.00	0.21	0.23	0.57	0.25	0.28	0.18	0.62	0.47	0.40	0.00	0.25	3.47	0.62
56	0.46	0.23	0.49	0.09	0.37	0.71	0.34	0.00	1.00	0.97	0.00	0.46	5.11	1.00
sum	19.05	12.49	17.21	19.16	15.00	18.87	13.00	14.31	17.31	21.71	15.24	13.72	197.07	

Normalised number of keystrokes taken by Group 2 to complete the non-trial tasks.

tasks/subjects	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	sum	max
8	11	9	9	18	9	9	9	16	9	9	9	9	126	18
9	34	38	4	42	33	36	36	15	35	37	118	35	463	118
11	14	14	14	14	23	14	14	32	14	14	14	14	195	32
12	3	3	13	3	41	14	11	22	12	11	11	11	155	41
14	12	12	14	12	12	12	12	35	12	12	16	17	178	35
16	17	15	17	15	19	15	15	24	15	15	17	17	201	24
18	17	12	9	9	9	65	16	14	16	12	19	9	207	65
20	12	14	12	14	12	12	12	15	12	12	18	12	157	18
21	71	26	26	131	24	27	77	19	25	234	31	17	708	234
22	7	7	7	3	7	7	4	15	7	7	7	7	85	15
23	37	17	11	114	12	13	27	23	21	96	30	12	413	114
24	39	29	24	24	14	39	24	33	25	32	28	24	335	39
25	13	44	14	22	18	90	12	22	14	18	24	27	318	90
26	43	33	21	20	19	8	22	18	19	8	70	10	291	70
27	3	19	24	15	10	106	48	53	10	75	12	11	386	106
28	36	37	57	73	24	41	303	42	47	39	95	67	861	303
30	48	17	19	17	17	23	20	63	21	22	19	17	303	63
31	9	47	30	26	25	17	37	26	15	18	24	22	296	47
32	18	21	15	68	51	9	12	16	57	84	12	11	374	84
33	18	24	19	17	17	17	17	18	18	18	21	16	220	24
34	47	67	14	13	12	20	27	75	14	11	13	47	360	75
35	22	8	26	58	11	6	33	48	9	8	4	22	255	58
36	6	10	7	6	9	19	6	20	3	4	7	6	103	20
37	54	15	13	14	20	14	2	15	17	20	20	13	217	54
38	6	6	5	5	6	5	5	6	7	6	6	5	68	7
39	6	8	7	7	4	7	9	14	9	7	7	9	94	14
40	38	15	21	13	13	13	13	26	17	13	15	11	208	38
41	11	19	10	10	26	12	80	14	14	31	15	22	264	80
42	11	7	7	5	5	9	13	16	5	31	7	7	123	31
43	48	85	52	19	65	41	85	37	19	20	26	58	555	85
44	6	10	11	14	9	14	32	11	33	19	20	11	190	33
45	8	7	1	6	1	7	2	1	7	8	6	1	55	8
46	1	1	1	1	1	7	5	8	7	1	1	1	35	8
47	1	1	1	1	1	1	4	7	1	1	1	1	21	7
48	7	17	6	6	7	7	7	14	6	5	9	6	97	17
50	9	13	10	9	10	10	9	15	10	9	10	9	123	15
51	28	58	67	18	24	18	67	28	23	30	33	36	430	67
52	8	13	6	22	12	12	21	14	9	18	12	9	156	22
53	15	23	17	8	16	16	18	17	10	45	21	10	216	45
54	45	145	194	41	9	14	194	32	194	122	116	21	1127	194
55	39	19	37	17	12	23	85	29	17	21	85	45	429	85
56	7	28	11	9	10	14	20	15	9	23	10	11	167	28
sum	885	1013	883	959	679	863	1465	983	844	1226	1039	726	11565	

Number of keystrokes taken by Group 1 to complete the non-trial tasks, after transformation to take success into account.

tasks/subjects	s13	s14	s15	s16	s17	s18	s19	20	s21	s22	s23	s24	sum	max
8	9	9	9	9	19	14	9	18	19	19	9	9	152	19
9	10	4	36	122	10	67	44	34	26	45	35	34	467	122
11	19	14	16	26	65	36	14	14	14	25	14	16	273	65
12	27	11	13	13	31	11	17	11	12	17	13	13	189	31
14	12	12	20	21	16	23	12	12	12	15	12	14	181	23
16	14	15	17	20	17	22	15	17	23	12	15	21	208	23
18	16	9	9	11	49	12	16	9	33	20	63	9	256	63
20	12	13	12	30	12	17	12	18	12	14	12	12	176	30
21	234	24	32	140	110	76	21	36	14	31	14	15	747	234
22	20	7	20	4	6	2	7	18	16	20	6	20	146	20
23	20	40	25	8	14	11	19	20	12	32	17	13	231	40
24	24	41	26	2	18	18	28	32	45	54	44	29	361	54
25	17	12	41	23	21	70	29	43	21	51	17	3	348	70
26	46	63	8	10	5	10	10	52	17	48	28	10	307	63
27	157	9	15	9	17	26	8	21	52	211	10	28	563	211
28	303	27	68	96	26	134	60	19	140	67	303	31	1274	303
30	63	19	35	17	39	23	17	24	33	56	63	29	418	63
31	64	14	35	16	20	47	10	11	29	43	21	24	334	64
32	12	8	108	20	20	64	39	37	54	210	16	11	599	210
33	17	19	22	16	15	27	8	18	21	21	19	22	225	27
34	14	23	64	75	74	21	23	51	15	75	14	27	476	75
35	58	8	4	10	4	6	52	7	4	29	4	4	190	58
36	39	11	6	2	18	3	8	3	5	3	14	3	115	39
37	16	10	25	14	44	44	14	16	14	11	14	13	235	44
38	6	8	8	7	6	6	7	8	6	7	6	6	81	8
39	11	8	11	8	7	4	8	8	14	12	7	4	102	14
40	38	13	18	21	11	38	16	11	13	18	38	11	246	38
41	16	9	145	151	10	9	24	39	16	126	10	9	564	151
42	8	14	42	6	5	8	11	18	8	56	5	12	193	56
43	85	14	68	19	47	25	29	85	23	85	55	52	587	85
44	10	6	85	15	14	31	6	65	41	14	14	13	314	85
45	2	3	51	12	1	17	6	8	51	7	9	6	173	51
46	5	1	2	7	1	8	1	1	3	8	1	3	41	8
47	6	1	1	8	1	8	1	1	1	3	1	5	37	8
48	14	7	6	6	7	6	16	10	6	17	7	6	108	17
50	9	14	17	17	9	14	11	17	12	14	17	9	160	17
51	67	19	19	33	18	27	28	67	22	67	67	19	453	67
52	36	23	19	15	12	33	9	11	6	36	10	18	228	36
53	27	16	31	13	17	39	13	46	32	32	13	17	296	46
54	194	50	90	112	69	194	30	27	194	194	194	16	1364	194
55	85	16	18	34	20	23	14	85	34	85	85	20	519	85
56	28	6	15	28	11	22	10	28	27	28	28	14	245	28
sum	1870	650	1312	1226	936	1296	732	1076	1152	1938	1344	650	14182	

Number of keystrokes taken by Group 2 to complete the non-trial tasks, after transformation to take success into account.

tasks/subjects	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	sum	max
8	0.58	0.47	0.47	0.95	0.47	0.47	0.47	0.84	0.47	0.47	0.47	0.47	6.63	0.95
9	0.28	0.31	0.03	0.34	0.27	0.30	0.30	0.12	0.29	0.30	0.97	0.29	3.80	0.97
11	0.22	0.22	0.22	0.22	0.35	0.22	0.22	0.49	0.22	0.22	0.22	0.22	3.00	0.49
12	0.07	0.07	0.32	0.07	1.00	0.34	0.27	0.54	0.29	0.27	0.27	0.27	3.78	1.00
14	0.34	0.34	0.40	0.34	0.34	0.34	0.34	1.00	0.34	0.34	0.46	0.49	5.09	1.00
16	0.71	0.63	0.71	0.63	0.79	0.63	0.63	1.00	0.63	0.63	0.71	0.71	8.38	1.00
18	0.26	0.18	0.14	0.14	0.14	1.00	0.25	0.22	0.25	0.18	0.29	0.14	3.18	1.00
20	0.40	0.47	0.40	0.47	0.40	0.40	0.40	0.50	0.40	0.40	0.60	0.40	5.23	0.60
21	0.30	0.11	0.11	0.56	0.10	0.12	0.33	0.08	0.11	1.00	0.13	0.07	3.03	1.00
22	0.35	0.35	0.35	0.15	0.35	0.35	0.20	0.75	0.35	0.35	0.35	0.35	4.25	0.75
23	0.32	0.15	0.10	1.00	0.11	0.11	0.24	0.20	0.18	0.84	0.26	0.11	3.62	1.00
24	0.72	0.54	0.44	0.44	0.26	0.72	0.44	0.61	0.46	0.59	0.52	0.44	6.20	0.72
25	0.14	0.49	0.16	0.24	0.20	1.00	0.13	0.24	0.16	0.20	0.27	0.30	3.53	1.00
26	0.61	0.47	0.30	0.29	0.27	0.11	0.31	0.26	0.27	0.11	1.00	0.14	4.16	1.00
27	0.01	0.09	0.11	0.07	0.05	0.50	0.23	0.25	0.05	0.36	0.06	0.05	1.83	0.50
28	0.12	0.12	0.19	0.24	0.08	0.14	1.00	0.14	0.16	0.13	0.31	0.22	2.84	1.00
30	0.76	0.27	0.30	0.27	0.27	0.37	0.32	1.00	0.33	0.35	0.30	0.27	4.81	1.00
31	0.14	0.73	0.47	0.41	0.39	0.27	0.58	0.41	0.23	0.28	0.38	0.34	4.63	0.73
32	0.09	0.10	0.07	0.32	0.24	0.04	0.06	0.08	0.27	0.40	0.06	0.05	1.78	0.40
33	0.67	0.89	0.70	0.63	0.63	0.63	0.63	0.67	0.67	0.67	0.78	0.59	8.15	0.89
34	0.63	0.89	0.19	0.17	0.16	0.27	0.36	1.00	0.19	0.15	0.17	0.63	4.80	1.00
35	0.38	0.14	0.45	1.00	0.19	0.10	0.57	0.83	0.16	0.14	0.07	0.38	4.40	1.00
36	0.15	0.26	0.18	0.15	0.23	0.49	0.15	0.51	0.08	0.10	0.18	0.15	2.64	0.51
37	1.00	0.28	0.24	0.26	0.37	0.26	0.04	0.28	0.31	0.37	0.37	0.24	4.02	1.00
38	0.75	0.75	0.63	0.63	0.75	0.63	0.63	0.75	0.88	0.75	0.75	0.63	8.50	0.88
39	0.43	0.57	0.50	0.50	0.29	0.50	0.64	1.00	0.64	0.50	0.50	0.64	6.71	1.00
40	1.00	0.39	0.55	0.34	0.34	0.34	0.34	0.68	0.45	0.34	0.39	0.29	5.47	1.00
41	0.07	0.13	0.07	0.07	0.17	0.08	0.53	0.09	0.09	0.21	0.10	0.15	1.75	0.53
42	0.20	0.13	0.13	0.09	0.09	0.16	0.23	0.29	0.09	0.55	0.13	0.13	2.20	0.55
43	0.56	1.00	0.61	0.22	0.76	0.48	1.00	0.44	0.22	0.24	0.31	0.68	6.53	1.00
44	0.07	0.12	0.13	0.16	0.11	0.16	0.38	0.13	0.39	0.22	0.24	0.13	2.24	0.39
45	0.16	0.14	0.02	0.12	0.02	0.14	0.04	0.02	0.14	0.16	0.12	0.02	1.08	0.16
46	0.13	0.13	0.13	0.13	0.13	0.88	0.63	1.00	0.88	0.13	0.13	0.13	4.38	1.00
47	0.13	0.13	0.13	0.13	0.13	0.13	0.50	0.88	0.13	0.13	0.13	0.13	2.63	0.88
48	0.41	1.00	0.35	0.35	0.41	0.41	0.41	0.82	0.35	0.29	0.53	0.35	5.71	1.00
50	0.53	0.76	0.59	0.53	0.59	0.59	0.53	0.88	0.59	0.53	0.59	0.53	7.24	0.88
51	0.42	0.87	1.00	0.27	0.36	0.27	1.00	0.42	0.34	0.45	0.49	0.54	6.42	1.00
52	0.22	0.36	0.17	0.61	0.33	0.33	0.58	0.39	0.25	0.50	0.33	0.25	4.33	0.61
53	0.33	0.50	0.37	0.17	0.35	0.35	0.39	0.37	0.22	0.98	0.46	0.22	4.70	0.98
54	0.23	0.75	1.00	0.21	0.05	0.07	1.00	0.16	1.00	0.63	0.60	0.11	5.81	1.00
55	0.46	0.22	0.44	0.20	0.14	0.27	1.00	0.34	0.20	0.25	1.00	0.53	5.05	1.00
56	0.25	1.00	0.39	0.32	0.36	0.50	0.71	0.54	0.32	0.82	0.36	0.39	5.96	1.00
sum	15.60	17.51	14.23	14.41	13.03	15.45	19.00	21.21	14.03	16.51	16.32	13.15	190.45	

Number of keystrokes taken by Group 1 to complete the non-trial tasks, after transformation to take success into account and normalisation.

tasks/subjects	s13	s14	s15	s16	s17	s18	s19	20	s21	s22	s23	s24	sum	max
8	0.47	0.47	0.47	0.47	1.00	0.74	0.47	0.95	1.00	1.00	0.47	0.47	8.00	1.00
9	0.08	0.03	0.30	1.00	0.08	0.55	0.36	0.28	0.21	0.37	0.29	0.28	3.83	1.00
11	0.29	0.22	0.25	0.40	1.00	0.55	0.22	0.22	0.22	0.38	0.22	0.25	4.20	1.00
12	0.66	0.27	0.32	0.32	0.76	0.27	0.41	0.27	0.29	0.41	0.32	0.32	4.61	0.76
14	0.34	0.34	0.57	0.60	0.46	0.66	0.34	0.34	0.34	0.43	0.34	0.40	5.17	0.66
16	0.58	0.63	0.71	0.83	0.71	0.92	0.63	0.71	0.96	0.50	0.63	0.88	8.67	0.96
18	0.25	0.14	0.14	0.17	0.75	0.18	0.25	0.14	0.51	0.31	0.97	0.14	3.94	0.97
20	0.40	0.43	0.40	1.00	0.40	0.57	0.40	0.60	0.40	0.47	0.40	0.40	5.87	1.00
21	1.00	0.10	0.14	0.60	0.47	0.32	0.09	0.15	0.06	0.13	0.06	0.06	3.19	1.00
22	1.00	0.35	1.00	0.20	0.30	0.10	0.35	0.90	0.80	1.00	0.30	1.00	7.30	1.00
23	0.18	0.35	0.22	0.07	0.12	0.10	0.17	0.18	0.11	0.28	0.15	0.11	2.03	0.35
24	0.44	0.76	0.48	0.04	0.33	0.33	0.52	0.59	0.83	1.00	0.81	0.54	6.69	1.00
25	0.19	0.13	0.46	0.26	0.23	0.78	0.32	0.48	0.23	0.57	0.19	0.03	3.87	0.78
26	0.66	0.90	0.11	0.14	0.07	0.14	0.14	0.74	0.24	0.69	0.40	0.14	4.39	0.90
27	0.74	0.04	0.07	0.04	0.08	0.12	0.04	0.10	0.25	1.00	0.05	0.13	2.67	1.00
28	1.00	0.09	0.22	0.32	0.09	0.44	0.20	0.06	0.46	0.22	1.00	0.10	4.20	1.00
30	1.00	0.30	0.56	0.27	0.62	0.37	0.27	0.38	0.52	0.89	1.00	0.46	6.63	1.00
31	1.00	0.22	0.55	0.25	0.31	0.73	0.16	0.17	0.45	0.67	0.33	0.38	5.22	1.00
32	0.06	0.04	0.51	0.10	0.10	0.30	0.19	0.18	0.26	1.00	0.08	0.05	2.85	1.00
33	0.63	0.70	0.81	0.59	0.56	1.00	0.30	0.67	0.78	0.78	0.70	0.81	8.33	1.00
34	0.19	0.31	0.85	1.00	0.99	0.28	0.31	0.68	0.20	1.00	0.19	0.36	6.35	1.00
35	1.00	0.14	0.07	0.17	0.07	0.10	0.90	0.12	0.07	0.50	0.07	0.07	3.28	1.00
36	1.00	0.28	0.15	0.05	0.46	0.08	0.21	0.08	0.13	0.08	0.36	0.08	2.95	1.00
37	0.30	0.19	0.46	0.26	0.81	0.81	0.26	0.30	0.26	0.20	0.26	0.24	4.35	0.81
38	0.75	1.00	1.00	0.88	0.75	0.75	0.88	1.00	0.75	0.88	0.75	0.75	10.13	1.00
39	0.79	0.57	0.79	0.57	0.50	0.29	0.57	0.57	1.00	0.86	0.50	0.29	7.29	1.00
40	1.00	0.34	0.47	0.55	0.29	1.00	0.42	0.29	0.34	0.47	1.00	0.29	6.47	1.00
41	0.11	0.06	0.96	1.00	0.07	0.06	0.16	0.26	0.11	0.83	0.07	0.06	3.74	1.00
42	0.14	0.25	0.75	0.11	0.09	0.14	0.20	0.32	0.14	1.00	0.09	0.21	3.45	1.00
43	1.00	0.16	0.80	0.22	0.55	0.29	0.34	1.00	0.27	1.00	0.65	0.61	6.91	1.00
44	0.12	0.07	1.00	0.18	0.16	0.36	0.07	0.76	0.48	0.16	0.16	0.15	3.69	1.00
45	0.04	0.06	1.00	0.24	0.02	0.33	0.12	0.16	1.00	0.14	0.18	0.12	3.39	1.00
46	0.63	0.13	0.25	0.88	0.13	1.00	0.13	0.13	0.38	1.00	0.13	0.38	5.13	1.00
47	0.75	0.13	0.13	1.00	0.13	1.00	0.13	0.13	0.13	0.38	0.13	0.63	4.63	1.00
48	0.82	0.41	0.35	0.35	0.41	0.35	0.94	0.59	0.35	1.00	0.41	0.35	6.35	1.00
50	0.53	0.82	1.00	1.00	0.53	0.82	0.65	1.00	0.71	0.82	1.00	0.53	9.41	1.00
51	1.00	0.28	0.28	0.49	0.27	0.40	0.42	1.00	0.33	1.00	1.00	0.28	6.76	1.00
52	1.00	0.64	0.53	0.42	0.33	0.92	0.25	0.31	0.17	1.00	0.28	0.50	6.33	1.00
53	0.59	0.35	0.67	0.28	0.37	0.85	0.28	1.00	0.70	0.70	0.28	0.37	6.43	1.00
54	1.00	0.26	0.46	0.58	0.36	1.00	0.15	0.14	1.00	1.00	1.00	0.08	7.03	1.00
55	1.00	0.19	0.21	0.40	0.24	0.27	0.16	1.00	0.40	1.00	1.00	0.24	6.11	1.00
56	1.00	0.21	0.54	1.00	0.39	0.79	0.36	1.00	0.96	1.00	1.00	0.50	8.75	1.00
sum	25.71	13.36	21.02	19.29	16.35	21.08	13.70	19.92	18.79	28.11	19.19	14.04	230.56	

Number of keystrokes taken by Group 2 to complete the non-trial tasks, after transformation to take success into account and normalisation.

tasks/subjects	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	sum
8	1	1	1	1	1	1	1	1	1	1	1	1	12
9	1	1	1	1	1	1	1	1	1	1	1	1	12
11	1	1	1	1	1	1	1	1	1	1	1	1	12
12	1	1	1	1	1	1	1	1	1	1	1	1	12
14	1	1	1	1	1	1	1	1	1	1	1	1	12
16	1	1	1	1	1	1	1	1	1	1	1	1	12
18	1	1	1	1	1	1	1	1	1	1	1	1	12
20	1	1	1	1	1	1	1	1	1	1	1	1	12
21	1	1	1	1	1	1	1	1	1	1	1	1	12
22	1	1	1	1	1	1	1	1	1	1	1	1	12
23	1	1	1	1	1	1	1	1	1	1	1	1	12
24	1	1	1	1	1	1	1	1	1	1	1	1	12
25	1	1	1	1	1	1	1	1	1	1	1	1	12
26	1	1	1	1	1	1	1	1	1	1	1	1	12
27	1	1	1	1	1	1	1	1	1	1	1	1	12
28	1	1	1	1	1	1	1	1	1	1	1	1	12
30	1	1	1	1	1	1	1	1	1	1	1	1	12
31	1	0	1	1	1	1	1	1	1	1	1	1	11
32	1	1	1	1	1	1	1	1	1	1	1	1	12
33	1	1	1	1	1	1	1	1	1	1	1	1	12
34	1	1	1	1	1	1	1	0	1	1	1	1	11
35	1	1	1	1	1	1	1	1	1	1	1	1	12
36	1	1	1	1	1	1	1	1	1	1	1	1	12
37	1	1	1	1	1	1	1	1	1	1	1	1	12
38	1	1	1	1	1	1	1	1	1	1	1	1	12
39	1	1	1	1	1	1	1	1	1	1	1	1	12
40	0	1	1	1	1	1	1	1	1	1	1	1	11
41	1	1	1	1	1	1	1	1	1	1	1	1	12
42	1	1	1	1	1	1	1	1	1	1	1	1	12
43	1	1	1	1	1	1	0	1	1	1	1	1	11
44	1	1	1	1	1	1	1	1	1	1	1	1	12
45	1	1	1	1	1	1	1	1	1	1	1	1	12
46	1	1	1	1	1	1	1	1	1	1	1	1	12
47	1	1	1	1	1	1	1	1	1	0	1	1	11
48	1	1	1	1	1	1	1	1	1	1	1	1	12
50	1	1	1	1	1	1	1	1	1	1	1	1	12
51	1	1	1	1	1	1	0	1	1	1	1	1	11
52	1	1	1	1	1	1	1	1	1	1	1	1	12
53	1	1	1	1	1	1	1	1	1	1	1	1	12
54	1	1	0	1	1	1	0	1	1	1	1	1	10
55	1	1	1	1	1	1	0	1	1	1	1	1	11
56	1	1	1	1	1	1	1	1	1	1	1	1	12
sum	41	41	41	42	42	42	38	41	42	41	42	42	495

Non-trial tasks completed successfully by Group 1.

tasks/subjects	s13	s14	s15	s16	s17	s18	s19	20	s21	s22	s23	s24	sum
8	1	1	1	1	0	1	1	1	0	1	1	1	10
9	1	1	1	1	1	1	1	1	1	1	1	1	12
11	1	1	1	1	1	1	1	1	1	1	1	1	12
12	1	1	1	1	1	1	1	1	1	1	1	1	12
14	1	1	1	1	1	1	1	1	1	1	1	1	12
16	1	1	1	1	1	1	1	1	1	1	1	1	12
18	1	1	1	1	1	1	1	1	1	1	1	1	12
20	1	1	1	1	1	1	1	1	1	1	1	1	12
21	0	1	1	1	1	1	1	1	1	1	1	1	11
22	0	1	0	1	1	1	1	1	1	0	1	1	9
23	1	1	1	1	1	1	1	1	1	1	1	1	12
24	1	1	1	1	1	1	1	1	1	1	1	1	12
25	1	1	1	1	1	1	1	1	1	1	1	1	12
26	1	1	1	1	1	1	1	1	1	1	1	1	12
27	1	1	1	1	1	1	1	1	1	1	1	1	12
28	0	1	1	1	1	1	1	1	1	1	0	1	10
30	0	1	1	1	1	1	1	1	1	1	0	1	10
31	1	1	1	1	1	1	1	1	1	1	1	1	12
32	1	1	1	1	1	1	1	1	1	1	1	1	12
33	1	1	1	1	1	1	1	1	1	1	1	1	12
34	1	1	1	0	1	1	1	1	1	1	1	1	11
35	0	1	1	1	1	1	1	1	1	1	1	1	11
36	1	1	1	1	1	1	1	1	1	1	1	1	12
37	1	1	1	1	0	1	1	1	1	1	1	1	11
38	1	1	1	1	1	1	1	0	1	1	1	1	11
39	1	1	1	1	1	1	1	1	1	1	1	1	12
40	0	1	1	1	1	1	1	1	1	1	0	1	10
41	1	1	1	1	1	1	1	1	1	1	1	1	12
42	1	1	1	1	1	1	1	1	1	1	1	1	12
43	0	1	1	1	1	1	1	0	1	0	1	1	9
44	1	1	1	1	1	1	1	1	1	1	1	1	12
45	1	1	0	1	1	1	1	1	1	1	1	1	11
46	1	1	1	1	1	1	1	1	1	1	1	1	12
47	1	1	1	1	1	1	1	1	1	1	1	1	12
48	1	1	1	1	1	1	1	1	1	0	1	1	11
50	1	1	0	1	1	1	1	0	1	1	0	1	9
51	0	1	1	1	1	1	1	0	1	0	0	1	8
52	1	1	1	1	1	1	1	1	1	0	1	1	11
53	1	1	1	1	1	1	1	1	1	1	1	1	12
54	0	1	1	1	1	0	1	1	0	0	0	1	7
55	0	1	1	1	1	1	1	0	1	0	0	1	8
56	0	1	1	0	1	1	1	0	1	0	0	1	7
sum	31	42	39	40	40	41	42	36	40	34	34	42	461

Non-trial tasks completed successfully by Group 2.

Appendix D

Content of The CD

The CD attached to this thesis provides the reader with the multimedia materials referenced in the thesis. This appendix lists the

D.1 Samples

Contains the audio samples references in this thesis. The samples are names after there reference in the thesis. For example, Sample 20 is called Sample20.wav. This directory contains 33 samples.

D.2 Experiment 1

This directory contains the sonified menu hierarchy used in the first experiment described in this thesis (*cf.* Chapter 4). The sonification is presented as a Java applet and can be viewed from a Java-enabled web browser. To launch the applet, double-click on the file 'Sonification.html'. The sounds used in this experiment are also available in the same directory as '.au' files.

D.3 Experiment 2

Simulation used for the experiment reported in Chapter 7. This directory does not only contain the simulation, it also includes all the data used for the experiment. The simulation will only run on a Windows32 platform. In order to launch to launch the simulation, double-click on the 'run.bat' file. Once the phone simulation appears, you can navigate in the menu using the top four navigation keys (see Chapter 7 for more details on the simulation). A sound should be played every time a menu selection is made. Make sure the volume is turned on!

This simulation also contains the complete experiment setup. Pressing the 'Return' key will bring up the instruction window. To view the next instruction, press the space bar. Note: the space bar only works if the telephone simulation window has he focus.

Here is a list and description of all the important files and directories present in this directory:

- run.bat

  This is the batch file that launches the simulation. To Launch the application, double-click on that file

- readme.txt

  A description of the content of the directory written for the on-line distribution of this application.

- instructions.txt

  List of instuctions displayed to the user at the beginning of the experiment. This file, as well as the following text files are used by the application to display the experiment instructions to the subjects.

- tasks.txt

  List of tasks that the subjects of the experiments were asked to complete.

- blind.txt

  Tasks that some subjects had to carry out with no visual feedback.

- debriefing.txt

  Questions asked to the subjects at the end of the experiment.

- soundlist.txt

  List of the 166 sounds used by the simulation, with their relative path.

- sounds

  Contains all the samples used by the simulation.

- jre

  Java Virtual Environment, required to run the simulation.

- Classes

  Java classes in which the simualtion is implemented.

D.4 AIDE

AIDE is the sonifiaction tool that has been presented in Chapter 8. This software only runs on Windows32 platforms. To run it, double-click on the 'run.bat' file. For more information on the software, please refer to Chapter 8. For more information on the functionality of AIDE, a small manual is available ('Manual.doc').

D.4.1 Using AIDE to listen to the sonification presented in Chapter 6

To listen to this sonification, select 'Open' in the 'File' menu. Then select 'Menu simulation'. This option will bring up a file chooser menu. Use this menu to choose the following file located on the CD: AIDE/Examples/8210 Sonification/simulation/final. Once this file is loaded, choose the 'Menu Sonification' tab of AIDE. You can now navigate in the menu and hear the sounds using the mobile telephone navigation keys. If you cannot hear any sound, ensure that you have selected a valid input and output in the MIDI menu.

Bibliography

[AE86] A. Aucella and S. F. Ehrlich. Voice messaging: Enhancing the user interface based on field performance. In *Proceedings of ACM CHI'86*, pages 156–161, Boston, 1986. ACM Press.

[Alt95] J. L. Alty. Can we use music in human-computer interaction? In *Proceedings of BCS HCI'95*, pages 409–423, Huddersfield, UK, 1995. Cambridge University Press.

[AV97] J. A. Alty and P. Vickers. The caitlin auralization system: hierarchical leitmotif design as a clue to program comprehension. In Elizabeth Mynatt and James A. Ballas, editors, *Proceedings of ICAD'97*, pages 89–95, Palo Alto, CA, USA, 1997. Xerox.

[AVR97] J. L. Alty, P. Vickers, and D. Rigas. Using musing as a communication medium. In *Proceedings of ACM CHI'97 Extended abstracts*, pages 30–31, Atlanta, GA, USA, 1997. ACM Press.

[BC71] S. A. Bregman and J. Campbell. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 89:244–249, 1971.

[BG92] M. Blattner and R. Greenberg. Communicating and learning through non-speech audio. In A.D.N. Edwards and S. Holland, editors, *Multimedia interface design in education*, pages 133–144. NATO ASI series, Speringer-Verlag, Berlin, 1992.

[BGB91] W. Buxton, W. Gaver, and S. Bly. Tutorial number 8: The use of non-speech audio at the interface. In *Proceedings of CHI'91*, New Orleans, 1991. ACM Press: Addison-Wesley.

[Big91] E. Bigand. La perception des schemas de tensions et detentes musicales dans une phrase musicale. *Analyse Musicale*, hors-serie:144–146, 1991.

[Bor42] E. G. Boring. *Sensation and perception in the history of experimental psychology.* New York, 1942.

[BP78] A. Bregman and S. Pinker. Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 32:19–31, 1978.

[Bre90] A. S. Bregman. *Auditory scene analysis: The perceptual organization of sound.* MIT Press, Cambridge, MA, 1990.

[Bre94] S. A. Brewster. *Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interaction.* Doctoral thesis, University of York, 1994.

[Bre97] S.A. Brewster. Navigating telephone-based interfaces with earcons. In H. Thimbleby, B. O'Conaill, and P. Thomas, editors, *Proceedings of BCS HCI'97*, pages 39–56, Bristol, UK, 1997. Springer.

[Bre98] S. A. Brewster. Using non-speech sounds to provide navigation cues. *ACM Transactions on Computer-Human Interaction*, 5(2):224–259, 1998.

[BRK96] S.A. Brewster, V.-P. Raty, and A. Kortekangas. Earcons as a method of providing navigational cues in a menu hierarchy. In A. Sasse, R. Cunnigham, and R. Winder, editors, *Proceedings of BCS HCI'96*, pages 169–183, London, UK, 1996. Springer.

[BRL91] W. Barfield, C. Rosenberg, and G. Levasseur. The use of icons, earcons and commands in the design of an online hierarchical menu. *IEEE Transactions on Professional Communication*, 34(2):101–108, 1991.

[BSG89] M. Blattner, D. Sumikawa, and R. Greenberg. Earcons and icons: Their structure and common design principles. *Human Computer Interaction*, 4(1):11–44, 1989.

[BW90] D. V. Beard and J. Q. Walker. Navigational techniques to improve the display of large two-dimensional spaces. *Behaviour and Information Technology*, 9(6):451–466, 1990.

[BWE95] S.A. Brewster, P.C. Wright, and A.D.N. Edwards. Experimentally derived guidelines for the creation of earcons. In M. Kirby, A. Dix, and J. Finlay, editors, *Adjunct Proceedings of BCS HCI'95*, pages 155–159, Huddersfield, UK, 1995.

[CB98] M. Crease and S. A. Brewster. Making progress with sounds - the design and evaluation of an audio progress bar. In *Proceedings of ICAD'98*, Glasgow, UK, 1998. British Computer Society.

[CBL94] S. Conversy and M. Beaudoin-Lafon. Monitoring background activities. In Gregory Kramer, editor, *Auditory Display*, pages 499–531. Addison-Wesley, 1994.

[Chi96] J. P. Chin. Personality trait attributions to voice mail user interfaces. In M. Tauber, editor, *Proceedings of ACM CHI'96 Conference Companion*, pages 248–249, Vancouver, Canada, 1996. ACM Press, Addison-Wesley.

[Cro93] R. G. Crowder. Auditory memory. In S. McAdams and E. Bigand, editors, *Thinking in Sound: The Cognitive Psychology Human Audition*. Oxford University Press, Oxford, 1993.

[CRS85] D. Canter, R. Rivers, and G. Storrs. Characterizing user navigation through complex data structure. *Behaviour and Information Technology*, 4:93–102, 1985.

[Cut76] J. E. Cutting. Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotomic listening. *Psycological Review*, 83:114–140, 1976.

[Dem89] L. Demany. Perception de la hauteur tonale. In M. C. Botte, G. Canvet, L. Demany, and C. Sorin, editors, *Psychoacoustique et Perception Auditive*, pages 43–82. Paris, 1989.

[Deu72] D. Deutsch. Octave generalization and tune recognition. *Perception and Psychophysics*, 11:411–412, 1972.

[DH86] W. J. Dowling and D. Harwood. *Music Cognition*. Academic Press, Orlando, 1986.

[DM90] G. Deffner and K. Melder. User acceptance and preference for advanced voice services features and dialogue styles. In *Proceedings of the Human Factors Society 34th Annual Meeting*, volume 1, pages 194–197, 1990.

[ECHP00] A. D. N. Edwards, B. P. Challis, J. C. K. Hankinson, and F. L. Pirie. Development of a standard test of musical ability for participants in auditory interface testing. In *Proceedings of ICAD'2000*, Atlanta, GA, 2000.

[EE47] H. W. Eagleson and O. W. Eagleson. Identification of musical instruments when heard directly and over a public address system. *Journal of the Acoustical Society of America*, 19:338–342, 1947.

[ER89] G. H. Engelbeck and T. Roberts. The effect of several voice-menu characteristics on menu selection performance. Technical report, US West Advanced Technologies, Englewood, Colo, 1989.

[ES90] T. Erickson and G. Salomon. Designing a desktop information system: observations and issues. In *Proceedings of ACM CHI'91 Human Factors in Computing Systems*, pages 49–54, New Orleans, 1990. ACM Press.

[Fle40] H. Fletcher. Auditory patterns. *Modern Physics*, 12:47–65, 1940.

[Fra58] R. Francès. *La Perception de la Musique*. Paris, second edition, 1972 edition, 1958.

[Fur86] G. W. Furnas. Generalized fisheye views. In *Proceedings of ACM CHI'86 Human Factors in Computing Systems*, volume 1, pages 14–18, Boston, 1986. ACM Press.

[Gav93a] W. Gaver. How do we hear the world? explorations in ecological acoustics. *Ecological Psychology*, 5(4):285–313, 1993.

[Gav93b] W. Gaver. What in the world do we hear? an ecological approach to auditory source perception. *Ecological Psychology*, 5(1):1–29, 1993.

[Gav97] W. Gaver. Auditory interfaces. In M. G. Helander, T. K. Landauer, and P. Prabhu, editors, *Handbook of human-computer interaction*. Elsevier Science, Amsterdam, second edition, 1997.

[GB84] J. D. Gould and S. L. Boies. Human factors challenges in creating a principal support office system - the speech filing system approach. *The ACM Transactions on Information Systems*, 1(4):273–298, 1984.

[Gib79] J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton-Mifflin, Boston, USA, 1979.

[Gre76] J. Grey. Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, 61(5):1270–1277, 1976.

[Gri96] N. Griffeth. Making a simple interface complex: Interactions among telephone features. In M. Tauber, editor, *Proceedings of ACM CHI'96 Conference Companion*, pages 244–245, Vancouver, Canada, 1996. ACM Press.

[GWM98] S. Goose, M. Wynblatt, and H. Mollenhauer. 1-800-hypertext: Browsing hypertext with a telephone. In *Proceedings of HyperText'98*, pages 287–288. ACM Press, 1998.

[Hel68] H. von Helmholtz. *Théorie physiologique de la musique.* Paris, 1868.

[HGH88] J. W. Hall, J. H. Grose, and M. P. Haggard. Comodulation masking release for multicomponent signals. *Journal of the Acoustical Society of America*, 83:677–686, 1988.

[HLJ+97] B. R. Huguenard, F. J. Lerch, B. W. Junker, R. J. Patz, and R. E. Kass. Working memory failure in phone-based interaction. *ACM Transactions on Computer-Human Interaction*, 4(2):67–102, 1997.

[HLML01] S. Helle, G. Leplâtre, J. Marila, and P. Laine. Menu sonification in a mobile phone - a prototype study. In *Proceedings of ICAD'2001*, Espoo, Finland, 2001.

[HN89] R. Halstead-Nussloch. The design of phone-based interfaces for consumers. In *Proceedings of ACM CHI'89 Conference on Human Factors in Computing Systems*, pages 347–352, Austin, Texas,USA, 1989. ACM Press, Addison-Wesley.

[How90] J. Howlett. May i speak to a person, please? *New York Times*, page 18, 1990, August 5 1990.

[How92] D. C. Howell. *Statistical Methods for Psychology.* PWS-KENT Pusblishing Company, third edition, 1992.

[How94] A. Howes. A model of the acquisition of menu knowledge by exploration. In *Proceedings of CHI'94*, pages 445–451, Boston, Massachusetts USA, 1994. ACM Press.

[HS88] S. Hart and L. Staveland. Development of nasa tlx (task load index): Result of empirical and theoretical research. *Human Mental Workload*, pages 139–183, 1988.

[IK93] P. Iverson and C. L. Krumhansl. Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America*, 94(5):2595–2603, 1993.

[Ins60] American National Standard Institute. Usa standard acoustical terminology. Technical report, American National Standards Institute, 1960.

[Ins92] European Telecommunications Standards Institute. Human factors (hf); specification of characteristics of telephone services tones when locally generated in terminals. Draft ETS 300 295 DE/HF-1003-B, European Telecommunications Standards Institute,, December 1992.

[JK96] B. E. John and D. E. Kieras. Using goms for user interface design and evaluation: Which technique? *ACM Transactions on Computer-Human Interaction*, 3:287–319, 1996.

[Kap95] V. Kaptelinin. A comparison of four navigation techniques in a 2d browsing task. In *Proceedings of the ACM CHI'95 Conference on Human Factors in Computing Systems*, volume 2, pages 282–283. ACM Press, 1995.

[Kig84] J. L. Kiger. The depth/breadth trade-off in the design of menu-driven user interfaces. *International Journal on Man-Machine Studies*, 20(2):201–213, 1984.

[Koh29] W. Kohler. *Gestalt Psychology.* Horace Liveright, New York, 1929.

[Kra94] Gregory Kramer. An introduction to auditory display. In Gregory Kramer, editor, *Auditory Display*, pages 1–77. Addison-Wesley, 1994.

[Kru64a] J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. *Psychometrika*, 29:1–27, 1964.

[Kru64b] J. B. Kruskal. Non-metric multidimensional scaling: a numerical method. *Psychometrika*, 29:115–129, 1964.

[Kru90] C. L. Krumhansl. *Cognitive Fundations of Musical Pitch.* Oxford University Press, Oxford, 1990.

[KZ89] D. Karis and B. Zeigler. Evaluation of mobile telecommunication systems. In *Proceedings of the Human Factors Society 33rd Annual Meeting*, pages 205–209, 1989.

[LB98] G. Leplâtre and S. A. Brewster. An investigation of using music to provide navigation cues. In *Proceedings of ICAD'98*, Glasgow, UK, 1998.

[LB00] G. Leplâtre and S. A. Brewster. Designing non-speech sounds to support navigation in mobile phone menus. In *Proceedings of ICAD'2000*, Atlanta, GA, 2000.

[LBH+97] S. K. Lodha, J. Beaham, T. Heppe, A. Joseph, and B. Zane-Ulman. Muse: A musical data sonification toolkit. In Elizabeth Mynatt and A. Ballas, James, editors, *Proceedings of ICAD'97*, pages 61–64, Palo Alto, CA, USA, 1997. Xerox.

[LJ83] F. Lerdahl and R.S. Jackendoff. *A Generative Theory of Tonal Music.* MIT Press, Cambridge, Mass, 1983.

[Mag96] M. Maguire. A human-factors study of telephone developments and convergence. *Contemporary Ergonomics*, pages 446–451, 1996.

[MBW98] E. Mynatt, M. Back, and R. Want. Designing audio aura. In *Proceedings of ACM CHI'98*, pages 566–573, Los Angeles, CA, USA, 1998. ACM Press.

[McA87] S. McAdams. Music: A science of the mind? *contemporary Music Review*, 2:1–61, 1987.

[McA97] S. McAdams. L'organisation perceptive de l'environnement sonore. In *Proceedings of the Rencontres IPSEN en ORL*, 1997.

[ME91] M. A. Miller and J. W. Elias. Using menus to access computers via phone-based interfaces. In *Proceedings of the Human Factors Society 35th Annual Meeting*, volume 1, pages 235–239, 1991.

[ME97] E. N. Mitsopoulos and A. D. N. Edwards. Auditory scene analysis as the basis for designing auditory widgets. In E. Mynatt and J. A. Ballas, editors, *Proceedings of ICAD'97*, pages 13–17, Palo Alto, CA, USA, 1997. Xerox.

[MF94] S. Mukherjea and J. D. Foley. Navigational builder: a tool for building navigational views of information spaces. In *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems*, volume 2, pages 289–209. ACM Press, 1994.

[MLL86] J. MacGregor, E. Lee, and N. Lam. Optimizing the structure of database menu indeses: A decision model of menu search. *Human Factors*, 28(4):387–399, 1986.

[Moo82] B. C. J. Moore. *An Introduction to the Psychology of Hearing*. Academic Press, London, 1982.

[MS96a] M. Marx and C. Schmandt. Clues: Dynamic personalized message filtering. In *Proceedings of ACM CSCW'96*, pages 113–121. ACM Press, 1996. -.

[MS96b] M. Marx and C. Schmandt. Mailcall: Message presentation and navigation in a nonvisual environment. In *Proceedings of ACM CHI'96 Conference on human factors in computing systems*, pages 165–172, Vancouver, BC Canada, 1996. ACM Press.

[Mur37] J.L. Mursell. *The Psychology of Music*. Norton, W.W., New York, 1937.

[MWD+95] S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff. Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychological Research*, 58:177–192, 1995.

[MWW90] M. M. Martin, B. H. Williges, and R. C. Williges. Improving the design of telephone-based information systems. In *Proceedings of the Human Factors Society 34th Annual Meeting*, volume 1, pages 198–202, 1990.

[Nor90] D. Norman. *The design of everyday things.* Currency Doubleday, New York, 1990.

[Nor91] K. L. Norman. *The psychology of menu selection: Designing cognitive control at the human/computer interface.* Ablex Publishing Corporation, 1991.

[Pat82] R. D. Patterson. Guidelines for auditory warning systems on civil aircraft. Technical Report CAA No. 82017, Civil Aviation Authority, 1982.

[Pel89] G. E. Pelton. Designing the telephone interface for voice processing applications. *Speech Technology*, 5(1):18–21, 1989.

[PHSV88] S. R. Parkinson, M. D. Hill, N. Sisson, and C. Viera. Effects of breadth, depth and number of responses on computer menu search performance. *International Journal of Man-Machine Studies*, 28:683–692, 1988.

[Pic89] J. O. Pickles. *Introduction to the Psychology of Hearing.* Academic Press, Cambridge, 1989.

[Plo76] R. Plomp. *Aspects of tone sensation.* Academic Press, London, 1976.

[PPS92] B. J. Pierce, S. R. Parkinson, and N. Sisson. Effects of semantics similarity, omission probability and number of alternatives on computer menu search. *International Journal on Man-Machine Studies*, 37(5):653–677, 1992.

[RA97] D. I. Rigas and J. L. Alty. The use of music in a graphical interface for the visually impaired. In S. Howard, J. Hammond, and G. Lindgaard, editors, *Proceedings of IFIP Interact'97*, pages 228–235, Sydney, Australia, 1997. Chapman & Hall.

[RAL97] D. I. Rigas, J. L. Alty, and F. W. Long. Can music support interfaces to complex databases? In *Proceedings of Euromicro-1997, New Frontiers of Information Technology*, Budapest, Hungary, 1997.

[Ras78] R. A. Rasch. The perception of simultaneous tones such as in polyphonic music. *Acustica*, 40:21–33, 1978.

[RE89] T. L. Roberts and G. Engelbeck. The effects of device technology on the usability of advanced telephone functions. In *Proceedings of ACM CHI'89 Conference on human factors in computing systems*, pages 331–337, Austin, Texas, USA, 1989. ACM Press.

[Roe75] Roederer. *Introduction to the Physics and Psychoacoustics of Music.* Springer-Verlag, New-York, second edition, 1975.

[Ros85] M. B. Rosson. Using synthetic speech for remote access to information. *Behaviour Research Methods, Instrumants and computers*, 17(2):250–252, 1985.

[RV92] P. Resnick and R. A. Virzi. Skip and scan: Cleaning up telephone interfaces. In *Proceedings of ACM CHI'92*, pages 419–426. ACM Press, 1992.

[RV95] P. Resnick and R. A. Virzi. Relief from the audio interface blues: Expanding the spectrum of menu, list, and form styles. *ACM Transactions on computer-human interaction*, 2(2):145–176, 1995.

[RW99] J. C. Risset and D. Wessel. Exploration of timbre by analysis and synthesis. In Diana Deutsch, editor, *The Psychology of Music*, Academic Press Series in Cognition and Perception. Academic Press, second edition edition, 1999.

[SASH93] L. J. Stifelman, B. Arons, C. Schmandt, and E. A. Hulteen. Voicenotes: A speech interface for a hand-held voice notetaker. In *Proceedings of ACM INTERCHI'93*, pages 179–186. ACM Press, 1993.

[SC62] E. L. Saldanha and John F. Corso. Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, 34:717, 1962.

[Sch40] M. Schoen. *The Psychology of Music.* Ronald Press, New York, 1940.

[Sch61] B. Scharf. Complex sounds and critical bands. *Psychological Bulletin*, 58:205–217, 1961.

[Sch64] B. Scharf. Partial masking. *Acoustica*, 14:16–23, 1964.

[Sch66] P. Schaeffer. *Traité des objets musicaux.* Paris, editions du seuil edition, 1966.

[Sch70] B. Scharf. Critical bands. In J. Tobias, editor, *Foundations of Modern Auditory Theory*. Academic Press, Orlando, 1970.

[Sch83] N. T. M. Scheffers. Simulation of the auditory pitch analysis: An elaboration on the dws pitch meter. *Journal of the Acoustical Society of America*, 74:1716–1725, 1983.

[Sch92] R. M. Schumacher. Phone-based interfaces: Research and guidelines. In *Proceedings of the Human Factors Society 36th Annual Meeting*, volume 2, pages 1051–1055, 1992.

[Sch93] C. Schmandt. Phoneshell: The telephone as computer terminal. In *Proceedings of ACM Conference on Multimedia'93*, pages 373–382. ACM Press, 1993.

[Sea38] C. Seashore. *Psychology of Music*. McGrawHill, New York, 1938.

[SH93] A. L. Schwartz and M. L. Hardzinski. Ameritech phone-based user interface standards and design guidelines. Technical report, Ameritech Services, Inc., April 12, 1993 1993.

[She62a] R. N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function. part i. *Psychometrika*, 27:125–140, 1962.

[She62b] R. N. Shepard. Multidimensional scaling with an unknown distance function. part ii. *Psychometrika*, 27:219–246, 1962.

[SHS95] R.M. Schumacher, M.L. Hardzinski, and A.L. Schwartz. Increasing the usability of interactive voice response systems. *Human Factors*, 37(2):251–264, 1995.

[SK91] P. A. Savage and D. G. Kemp. Auditory versus visual presentation of help messages. In *Proceedings of the Human Factors Society 35th Annual Meeting*, volume 1, pages 244–248, 1991.

[Sla85] W. Slawson. *Sound Color*. University of California Press, 1985.

[Slo86] J. Sloboda. Cognition and real music: The psychology of music comes to age. *Psychologica Belgica*, 26:199–219, 1986.

[Smo94] S. W. Smoliar. The role of music in multimedia. *IEEE Multimedia*, pages 9–11, 1994.

[Spe80] N. Spender. Psychology of music (i-iii). *The New Grove's Dictionary of Music and Musicians*, 15:388–427, 1980.

[SS93] A. L. Schwartz and E. C. Schwab. Improving menu design for the rapid order audiotext system using cluster analysis. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, volume 1, pages 230–234, 1993.

[Vir91] R. A. Virzi. A preference evaluation of three dialing plans for a residential, phone-based, information service. In *Proceedings of the Human Factors Society 35th Annual Meeting*, volume 1, pages 240–243, 1991.

[Viv73] O. Viver. *Varèse*. Editions du Seuil, Paris, France, 1973.

[VN75] L. P. A. S. Van Noorden. *Temporal Coherence in the Perception of Tone Sequences*. Phd thesis, Eindhoven University, 1975.

[War92] K. Warren. Helping the desk with a telephone menu. In *Proceedings of ACM SIGUCCS'92 Conference on User Services*, pages 251–252. ACM Press, 1992.

[WB99] A. Walker and S. Brewster. Trading space for time in interface design. In *Proceedings of Interact'99*, volume 2, pages 67–68, Edinburgh, UK, 1999. British Computer Society.

[WC98a] D. Williams and C. Cheepen. "just speak naturally": Designing for naturalness in automated spoken dialogues. In *Proceedings of CHI'98*, volume 2, pages 243–244. ACM Press, 1998.

[WC98b] D. Williams and C. Cheepen. "the sound of silence": A preliminary experiment investigating non-verbal auditory representations in telephone-based interfaces automated spoken dialogues. In *Proceedings of ICAD'98*, Glasgow, UK, 1998.

[Wen35] E. C. Wente. Characteristics of sound transmission in rooms. *Journal of the Acoustical Society of America*, 7:123, 1935.

[Wer38] M. Wertheimer. *Source Book of Gestalt Psychology*. Hartcourt, Brace and Co, New York, 1938.

[Wes79] D. L. Wessel. Timbre space as a musical control structure. *Computer Music Journal*, 3(2):45–52, 1979.

[WG72] L. Wedin and G. Goude. Dimension analysis of the perception of instrument timbre. *Scandinavian Journal of Psychology*, 13:228–240, 1972.

[Wis94] T. Wishart. *Audible Design*. Orpheus the Pantomine Ltd, 1994.

[WKK95] C. Wolf, L. Koved, and E. Kunzinger. Ubiquitous mail: Speech and graphical interfaces to an integrated voice/email mailbox. In *Proceedings of Interact'95*, pages 247–252, Lillehammer, Norway, 1995. Chapman and Hall.

[Yan94] N. Yankelovich. Speechacts and the design of speech interfaces. In *Proceedings of CHI'94 Conference on Human Factors in Computing Systems*, Boston, MA, 1994. ACM Press.

[Yan97] N. Yankelovich. Using natural dialogues as the basis for speech interface design. In *Automated Spoken Dialog Systems*. MIT Press, 1997.

[YLM95] Nicole Yankelovich, G. Levow, and M. Marx. Designing speechacts: Issue in speech and user interfaces. In *Proceedings of CHI'95 Conference on Human Factors in Computing Systems*, Denver, CO, 1995. ACM Press.

[ZF90] E. Zwicker and H. Fash. *Psychoacoustics Facts and Models*. Springer-Verlag, Berlin, 1990.