Principles for Improving Interaction in Telephone-Based Interfaces

Stephen A. Brewster
Department of Computing Science, University of Glasgow

INTRODUCTION

Telephone-based interfaces (TBI's) are becoming an increasingly important method for interacting with computer systems. The telephone is an ubiquitous device and is many people's primary method of entry into the information infrastructure. Access to an increasing number of services is being offered over the telephone, such as voice-mail and electronic banking. The rapidly increasing use of mobile telephones means that people access these services at many different times and places. Telephones themselves are now also incorporating greater functionality (such as multi-party calling or call forwarding). The provision of this extra functionality may be rendered useless if usability issues are not considered [24]. This project will improve the design of TBI's and will be a major benefit for the UK telecommunications industry.

The telephone itself allows only a limited form of interaction. There is no graphical display (although some do have small textual displays) so output is limited to recorded or synthesised speech and simple sounds. Users provide input via the keypad (although speech recognition is sometimes used). These techniques are limited and reduce the usability of telephone-based systems [28].

One common problem when interacting in TBI's is that users get lost in the hierarchy of menus that they must go through to reach an option or function [27, 30]. For example, in a telephone banking system users might call their bank and navigate through a hierarchy of voice menus to find the service required. However, they may get lost in the menu structure before they get the option they want because little navigation information is presented. Yankelovich *et al.* [31] say: "These [telephone-based] interfaces, however, are often characterised by a labyrinth of invisible and tedious hierarchies which result when menu options outnumber telephone keys or when choices overload users' short-term memory". One reason for this is that speech alone cannot provide sufficient feedback. As Wolf *et al.* [30] say: "The biggest problems participants had with the speech interface were trying to find specific messages, getting messages quickly and knowing where they were in the mailbox". Speech is serial and slow, this limits interactions so little feedback is given about current location. The more navigation information that is given in speech, the longer it takes and the more it obstructs the information the user is trying to access. We suggest that non-speech sound can solve many of these problems.

There are guidelines for the design of TBI's [24, 28] but these include nothing about the use of non-speech sound. **Therefore, the innovative aspect of this proposal is to use structured non-speech sound to enhance the output of information in TBI's.** There is a growing body of research which indicates that the addition of non-speech sounds to human-computer interfaces can improve performance and increase usability [2, 4, 6, 22]. My previous work [12, 15] showed that adding sound to graphical interfaces gave significant qualitative and quantitative improvements. This is because sound has many advantages. For example, it is good for communicating information quickly [9]. Unlike speech, non-speech sound is universal; the user is not tied to one language, which is important for European markets. There is also great potential for the results of this work in other non-graphical interfaces such as those for visually disabled people and those where working conditions or protective clothing mean that a screen cannot be used. We will ensure that our results can also be used in these areas.

We will investigate the best ways of combining speech and non-speech sounds to create multimedia TBI's. The UK Technology Foresight Programme identified multimedia and sound as two of the specific technology opportunities available to the UK. It also made clear that user needs and ease of interaction were priorities. The research proposed here will study this and investigate the possibilities for improving interaction using sound.

In addition to navigation, non-speech sounds can be used to provide a richer form of feedback for new interaction techniques in TBI's. Therefore the other aspect of this project is to investigate new interaction techniques (which are currently impossible due to the poor feedback available).

My own work provides solid foundations for this project. Initial experiments showed that sounds can give navigation cues [1, 8, 9, 11]. This unique approach has attracted international attention from industry (I am running student projects with IBM and Nokia in this area and Telecom Sciences Corporation Ltd. will participate in this project) and academia. This work has also been published at one of the major Human-Computer Interaction conferences [9].

I am initially applying for pump-priming funding as this is my first EPSRC proposal. I am requesting one research student for the project. Our progress will be measured against the following milestones:

- First year: Develop principles, based on experimental results, for using structured sounds as navigation cues. Produce report.
- Second year: Full test of navigation cues in real-world system. With the help of our industrial partners, ensure that the cues are effective, simple to learn and simple to use. Produce guidelines report. Produce designs for new interaction techniques using sound.
- Third year: Demonstrate effectiveness of new techniques in full real-world evaluation of project. Produce final report. Host workshop for industry and academia to disseminate knowledge gained.

BACKGROUND

The use of structured non-speech sounds at the human-computer interface is a new research field. In the short time it has been under investigation some significant benefits have been identified [6, 18, 22]. Much of this work has focused on the use of sound in graphical interfaces [6]. There is also great potential for such sounds in interfaces where there is no graphical feedback, such as telephone-based interfaces, interfaces for visually disabled people and interfaces where working conditions or protective clothing mean that a screen cannot be used. This project will focus on TBI's but the results will provide insight into the problems faced in these other domains. The research is novel because currently no other group is systematically studying the use of non-speech sound to improve telephone-based human-computer interfaces. This proposal, however, will build upon the results of researchers working in related areas.

Rosson [27] did some of the earliest work in the design of telephone-based interfaces. She noticed that the problem of navigation caused users great difficulty. She investigated a hierarchical telephone-based interface for travel/visitor information. A user could call the system and move through the hierarchy of voice menus to find information such as addresses and phone numbers of restaurants. She describes one common problem: "It is important to note that the information needed to convey position in the hierarchy was implicit in the content of the utterances users heard". Feedback confirming that one had moved from the top to the middle level of the hierarchy was available only by understanding a category/sub-category relationship. After hearing "Restaurants" and making a 'Down' move, the user might hear "Chinese" and would have to infer that a move to a lower level of the hierarchy had been made. She suggested that this may have been the source of many of the users problems.

Rosson suggested that one way to solve the problem was to give extra speech feedback. For example "You have moved to the next item in the Chinese Restaurant list. It is…". She suggested, however, that this would make the interface appear slow and clumsy. This extra feedback might also be longer than the information being retrieved and so obscure it. Such feedback was rejected for this very reason by Stevens and colleagues [29] when designing navigation cues in a system to provide non-visual access to mathematics (see below). Rosson suggested other methods: "More attractive possibilities are to increase the information implicit in the utterance itself, by systematically varying the syntax of the utterances at different levels, or by assigning a different "voice" to each level".

There are problems with both of these methods. Varying the syntax might result in complex messages that again obscure the information being communicated. The relatively low quality of sound over telephone equipment may also reduce the listener's ability to differentiate many different voices (for example, you do not always recognise acquaintances when they telephone you). Speech is also slow so the navigation information would get in the way of the information the user was trying to find.

We suggest an alternative solution: Structured non-speech audio messages. These can provide a hierarchical system of sounds that could represent a menu hierarchy. Meera Blattner and colleagues from the Lawrence Livermore National Laboratory initially developed the idea of structured audio messages called *earcons* [4, 5]. These are abstract, musical sounds built up from smaller units that can be manipulated to build complex structures. They proposed methods for constructing earcons but never evaluated them. As part of my PhD research I performed the first detailed investigation of earcons [6, 14]. These sounds will constitute the auditory feedback used in the project. Professor Blattner and I are currently working together on a book about non-speech audio.

Figure 1 shows a simple hierarchy of earcons based on one possible family of errors. Each earcon is a node on a tree and inherits the properties of the earcon above it. The different levels are created by manipulating the parameters of earcons (for example, rhythm, pitch, timbre). In the diagram the top level of the tree is a neutral earcon representing a family of errors. It has a flute timbre. The structure of the earcon from level one is inherited by level two and then changed. At level two there is still a continuous flute sound but new timbres are added to play alongside it. At level three a rhythm is added to the earcon from level two to create a sound for a particular error. This rhythm is based on the timbre from the level above. In the case of the overflow error there would be a continuous flute sound with a three note rhythm played on an organ accompanying it.

Using earcons, this hierarchy is easily extensible. For example, to add another major category of errors all that is needed is a new timbre. To create a new type of execution error only a new rhythm is needed and it can be added to the existing hierarchy. Therefore earcons provide a very flexible system for representing telephone-based menu structures. The structure could be extended and users would not require retraining. In such a system the sounds would play continuously (but quietly) in the background at each level, giving location information. Users could listen to the current sound and from it work out their location in the hierarchy. The sounds would make explicit the differences moving from level to level or across the same level because the sounds would be related in different ways. Wolf *et al.* [30] confirm the usefulness of this approach and suggest that this might be a solution to some of the problems they had with their combined voicemail and email speech interface system: "Replacing much of the text-to-speech feedback with brief distinctive earcons would make traversal of the mailbox more efficient".

One important area of research that will provide valuable input into this project is the design of interfaces for blind people. Alistair Edwards at the University of York has done key work in the field of improving interfaces for blind and visually-impaired people [19, 20]. In recent research, I worked with him and his colleagues to build a system for presenting mathematics to blind people [29]. It used earcons to provide a 'glance' at the overall structure of an algebraic expression. This gave some idea of the complexity of the expression before users started to browse it with synthetic speech. The glance allowed users

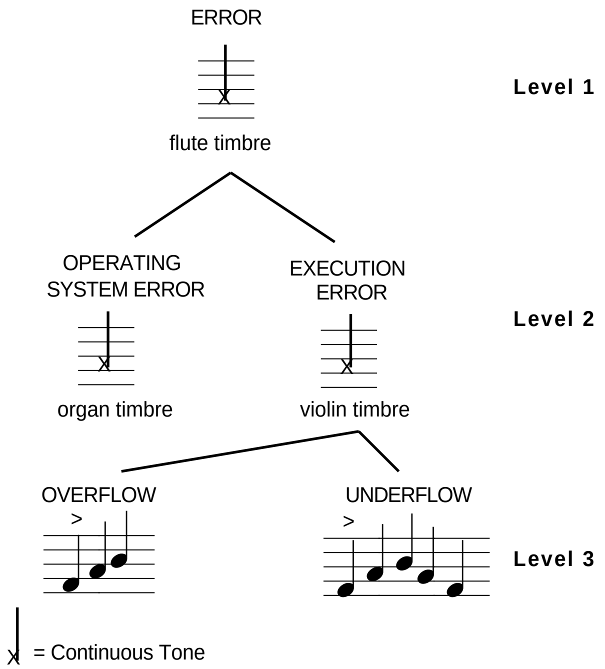

Figure 1: A hierarchy of earcons representing errors.

to choose the appropriate browsing strategy so that they were not overwhelmed with synthetic speech. They did this without interfering with the synthetic voice presenting the mathematics. The cues were also much shorter than an equivalent voice message. Many of the principles used for the design of systems for blind people can be applied to the design of the telephone-based interfaces.

Bill Gaver, Royal College of Art, has developed an alternative to earcons for presenting information in sound, called *auditory icons*. For example, he used these to improve the interface for a process control system [22]. Auditory icons are based on everyday sounds that have an intuitive link to an operation or action in the interface. Some of the strengths of auditory icons (such as their naturalness) will be used in the design of the sounds for this project.

There is much work in the area of alarms and warnings, but as yet little of this has been applied to user-interface design. In particular, Roy Patterson [25] and Judy Edworthy [21] have put forward guidelines for the design of alarms. These include ways of controlling the annoyance and urgency of auditory feedback. These guidelines will be applied to the design of the sounds used for the project.

I have also done work in this area providing solid foundations for this project. I have performed the most in-depth studies of the use of earcons [6, 13]. I have also produced a set of guidelines for designers to enable them to create effective earcons [16]. These guidelines were an initial attempt at making the use of sound easier, but they were only the first step. They helped in the design of the earcons themselves but not in how to use them to improve interaction.

I and colleagues have experimentally tested the use of earcons to provide navigation cues in simple menu hierarchies [1, 8, 9, 11]. This initial research has been very promising. Participants were able to work out where earcons would locate them in the hierarchy with over 80% accuracy, indicating that they were able to use the earcons to work out their position in the hierarchy. We also tested the ability of participants to recognise new sounds based on the ones they had already heard. Again their ability to do this was good with over 80% recognition rates. This indicated that earcons could provide good navigation cues and also were extensible. This technique also makes possible the use of hyperlinks in purely auditory interfaces. This unique approach has attracted international attention and has been published at one of the major human-computer interaction conferences [9]. It shows that these ideas have been assessed and seen as effective.

In recent research, I and one of my students [1] built upon this navigation work. In this second experiment we tested other aspects of the earcons. For example, we found no difference in performance between the CD quality sounds used in the first experiment and the lower telephone quality sounds we used in this experiment. We also found that listeners could recall the earcons well over time. This is important because users may not use a telephone service frequently but would need to remember the sounds. We also began to look at some of the different training techniques that could be used to teach users the sounds. These initial experiments show the approach has potential. In this project we hope to demonstrate unequivocally that this potential can be realised.

The other main strand of my work has been the use of sound to improve the usability of graphical human-computer interfaces. I and colleagues have experimentally tested the use of earcons in buttons and scrollbars and found that, in both cases, usability was significantly improved: Both the time taken to recover from errors and the time taken to complete tasks were reduced [7, 12, 15]. The research in the project described here and my work on graphical user interfaces will feed into each other; any increased knowledge about the creation and use of earcons will improve both aspects of my research.

PROGRAMME AND METHODOLOGY

The overall aim of this project is the creation of more usable TBI's. **The major innovative aspect is the use of structured non-speech audio messages to convey complex information to users.** With the rapid increase in the use of mobile telephones and the similarly rapid increase in the number of telephone services offered it is important that TBI's are improved. The current technology of speech output and keypad input has changed little over the last ten years. However, to get access to more complex data a more powerful user interface is needed. This project will investigate the use of non-speech sounds to provide navigation cues in a menu hierarchy and will also look at combining sounds and new interaction techniques to improve usability.

One of the main problems with TBI's is the narrow communication channel between the computer and the user. Output is usually restricted to recorded or synthetic speech. This narrow bandwidth of communication makes telephone-based systems hard to use.

There are currently no other research groups investigating the use of non-speech sound to improve TBI's. However, my initial experiments have shown enough success to indicate that further, more detailed investigation is necessary [1, 8, 9, 11]. This work has also attracted the attention of industry. Currently I am running student projects based on this work with IBM and Nokia. Telecom Sciences have also agreed to be an industrial partner in this project donating equipment and resources for us to build and test our ideas. Having industrial partners in this research will allow us to gather requirements from them and their customers. We will also be able to test our designs in real-world environments.

As described above, there are many experimentally-proven advantages to using sound at the interface. However, most previous examples were *ad hoc* solutions because the use of sound to improve human-computer interfaces is a new area and therefore there has been little to build upon. One major problem holding up the acceptance and use of sound by industry and academia is that there are no guidelines to aid a designer in how to use sound effectively [3]. We will therefore undertake the first systematic evaluation to find out how sound can be used to improve usability.

There are four main stages to the work described in the proposal: 1. Extend my initial research into the use of sound to provide navigation cues; 2. Develop a generalised method that designers can follow to allow them to add audio navigation cues to their TBI's; 3. Use sound to provide more sophisticated output from new interaction techniques; 4. Evaluate the work on a real-world TBI. Reports will be produced after each of these stages and can be used as a milestones to measure our progress through the project.

First year

The first year will focus on the extension of my initial research into using earcons to provide navigation cues in TBI's. This work showed that earcons were a good way of providing navigation cues. Participants reached a high level of recall after only a short period of training. They could also recognise new earcons based on the rules they had learned. However, there were problems with the mapping used, which mapped musical parameters (instrument, pitch, rhythm, etc.) to the level of the hierarchy. Representing large or irregular hierarchies was difficult because not enough parameters are available. The Research Student (RS) and I would investigate new methods for representing hierarchies. We have the data from my initial experiments as a base against which to compare any new techniques

For this aspect of the research the RS would investigate alternative methods of using earcons for navigation cues. He/she will try different mappings of sound parameters to nodes in a menu hierarchy to try and find a method which gives high rates of recall, is simple to learn and simple for designers to construct. One potentially powerful method which I have developed but not yet tested is the idea of basing the hierarchies around a decimal system. See Figure 2 for an example. Using this method a set of simple earcons would be created for 0 to 9 and dot. These simple sounds would then be combined to create a hierarchy as shown in the figure. This method is more

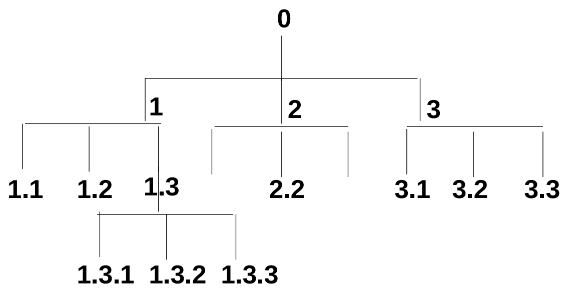

Figure 2: A decimal representation of a hierarchy

general than the one used in my previous experiments and it would allow hierarchies of arbitrary depths to be created. However, it has not yet been evaluated to determine its effectiveness.

These different techniques will be experimentally tested to discover which is the most effective. The experimental methods used will be based on my previous work [9]. I have considerable experience in designing and conducting such experiments so I will be able to help the RS design the experiments quickly and effectively. Each investigation will take approximately three months (based on the time taken for previous experiments).

The study of a particular set of earcons will take the following form. The earcons will be developed to represent a hierarchy. We will use the reference hierarchy I have tested in the past [9] to give a us a baseline of performance to compare the different sets of sounds we create. The sounds will be constructed around the guidelines for earcon construction by Brewster *et al.* [16].

The experiments will be of a standard form. Participants will be given initial training on the earcons they are going to hear. They will then have five minutes to listen to the sounds themselves (this training is similar to that which a new subscriber to a telephone system might get. He/she would get instructions on the system from the service provider, then perhaps get five minutes of free call time to try out the system). For testing the participants will hear a sound and then have to indicate where it fits into the hierarchy. We will also investigate the participants' abilities to recognise new earcons that they have not heard before (but based on the same rules used to construct the rest of the earcons).

This design has been successful in testing previous sets of earcons. Quantitative measures of recall rates will be taken. These will be accompanied by qualitative measures of subjective workload, overall preference and annoyance. Combining quantitative and qualitative measures gives a full measure of the usability of a system. One of the main concerns potential users of auditory interfaces have is annoyance. This will be explicitly measured to make sure it is not a problem. Results from the experiments will be statistically tested to compare them against the reference set. We will produce a report describing the results from the experiments completed.

The design of the sounds and the design of the experiments will give the RS a good background necessary for he skills he/she will need for the rest of the PhD. I will provide guidance on the experiments but the

RS will have the flexibility to use his/her own creativity to find novel solutions. The RS will be given the opportunity to investigate other issues that he/she finds interesting, as long as they fall within the boundaries of the project. As this research is the first of its kind there are many unanswered questions. For example, one possibility provide by this research (as mentioned above) is the use of hyperlinks within a telephone-based system. This needs further investigation to find out the best ways of using such links. The benefits they offer in graphical interfaces indicate that they may have potential in non-visual interfaces. Training techniques are another area where more investigation is necessary. Previous work [1] showed that different techniques have significant effects on recall rates. This needs more formal investigation.

Second year

The RS will finish any remaining experiments on the navigation earcons. We will then have a set of earcons that can provide easily learned and recalled non-speech navigation cues. We will then test these cues in a real TBI with our industrial partners. In this experiment we will present the same TBI to two different groups of users. One group will use the standard system with no navigation cues, the other will use the enhanced system. We will set the users typical tasks to perform and measure their performance. Our industrial partners will supply us with the TBI platform and will be able to help us develop a typical set of tasks. Using the platform as a test-bed it will be straightforward to create example TBI's.

We will measure the participants' performance in terms of time and error rates. Using the system with auditory navigation cues participants' should be able to complete the tasks more quickly (the navigation cues will help them navigate faster) and with fewer errors (the navigation cues will help them get to the menu option they require more accurately). As above, we will not only measure objective data. We will look at participants subjective workload when using the system, along with preference and annoyance. We hope that the sounds should lower workload as participants will find it easier to use the system. They should also not find the system more annoying to use and therefore prefer it to the standard system. Results from this experiment should give us strong data about the ability of sounds to provide navigation cues in a real world task. A paper will be written on the results.

From this work a set of guidelines for creating navigation earcons will be developed by the RS and myself. The guidelines will take a TBI designer through the stages necessary to design the hierarchy of sounds needed for his/her system and to create the sounds to represent it. This will be written up as a paper. These guidelines will be integrated into existing guidelines for TBI's [24, 28] A set of sounds will be made available over the Internet that a designer will be able to use along with the guidelines. This will simplify development for designers because they will not have to create all of the sounds from scratch. Many interface designers are not sound experts so we will simplify their interface development by providing a set of guidelines that mean they can integrate the sounds we provide into their systems, safe in the knowledge that they will be improving the usability of their TBI's.

I will then evaluate the guidelines with our industrial partner. This will enable us to ensure that the guidelines themselves are usable by TBI designers in industry. This will take the form of a walkthrough of the guidelines with the designers. Any problems can then be identified and corrected. The final set of guidelines will be made available over the Internet and published as a paper.

The next stage of the work for the RS will be to begin to design new interaction techniques based on recent research and the rich feedback provided by earcons. This rich feedback will not only provide navigation cues but will also allow greater feedback from interacting with the TBI. The RS will have gained the necessary skills in earcon design and have seen many of the interaction problems.

New techniques such as 'skip and scan' [26], alphawheels [23] and speech recognition [31] improve the range of input techniques available to users of TBI's. However, these input techniques themselves can be improved with more sophisticated feedback. This will allow the system to indicate to the user the effects of their actions without getting in the way of any speech which is giving the users information.

These new interaction techniques will be evaluated as before. Experimental evaluations will take place to discover if the new techniques improve usability. We will again use quantitative and qualitative measures to investigate our new designs. Any problems that are found will be analysed and the solutions incorporated into a new version of the technique.

Third year

The final year of the project will involve the RS finishing off any remaining experiments into the new interaction techniques developed. The results from these experiments will be published as a report. We will then begin a final evaluation of the work that has been done. A real-world system will be developed based on a complex TBI. Again this will be based on the TBI platform from our industrial partners. It will incorporate our navigation earcons and our new interaction techniques. This system will be an exemplar of what can be done with sounds in TBI's. It will show what will be possible for such interfaces in the future.

This final evaluation will allow us to test the interaction techniques in a realistic context and also in combination with the navigation sounds. We will use similar evaluation methods to our evaluations described above. Our industrial partner will have strong input in creating this system and the tasks the participants will perform. We will measure time to complete tasks, error rates, workload, annoyance and user preference. Results will show the advantages gained from richer feedback to the user.

The results of this full-scale evaluation will allow us to refine our ideas further. Any problems that are found will be corrected and, if necessary fixed and re-tested. We will then incorporate the new evaluation techniques into the guidelines for the design of TBI's. These will then allow TBI designers to make use of this research.

As the final part of this project we will hold a workshop to allow us to communicate our findings to academia and industry. The work will then all be brought together and published as a final report.

RELEVANCE TO BENEFICIARIES

There are three main beneficiaries of this research:

Telephone service developers will benefit because it will be easier for them to create usable information services. They will be able to include auditory cues that will allow users to navigate around their interfaces. They will also be able to make use of more flexible interaction techniques because of a richer set of available output techniques. They will be able to create new TBI's that will enable the use of more complex and useful information over the telephone.

End-users of telephone services will benefit because of the increased richness of their interaction with telephone-based systems. Current systems will be made easier to use and to navigate around. Future interfaces will be created that allow access to more complex (and useful) data than is now possible.

Blind computer users will benefit because, just as work from the area of interface design for blind people will feed into this work, so this work will feed back into interface design for blind people. Currently there is little use of non-speech audio in interfaces for blind people. These interfaces mainly make use of synthetic speech and suffer from many of the problems of feedback in TBI's. Blind users often need to navigate through hierarchies of information, for example file system hierarchies, and currently there is no good method of doing this. Results useful in this area will be published in the appropriate journals.

This work will also enhance our overall knowledge of earcons and what they can do. This will feed into research in other areas, for example sonically-enhanced graphical interfaces.

DISSEMINATION AND EXPLOITATION

Papers on the results of each stage of the work will be presented at national and international conferences so that industry and the academic community can see the work. Conferences will be targeted that attract delegates from both academia and industry such as BCS HCI, ACM CHI and ICAD. The work will also be published in a more detailed form in journals.

The guidelines for TBI designers will be made available on the World Wide Web in electronic form so that it will be freely available to all who want to read them. It is particularly important that this work is presented in an interactive form because the sounds cannot easily be understood by reading about them. Therefore the Web is a useful dissemination tool as we can provide interactive demonstrations, including the sounds with the papers that discuss them.

We will provide a set of sounds that interface designers can use in their systems. These will be in standard formats and therefore usable in many different computing environments and will be download-able over the Internet. Other sounds that we create will be reserved for future commercial exploitation by UK industry.

I will host a workshop in the last year of the project to communicate our ideas to academia and industry. This will allow us to demonstrate our ideas for the future of TBI's. A final report will be produced that brings together all of the knowledge gained.

JUSTIFICATION OF RESOURCES

As this is the first EPSRC project I have applied for, I am requesting some general resources and some for this particular project. Our industrial partners will supply us with a TBI development platform and development tools. The following resources are requested:

One Research Student is required to do the design and experimental testing of the enhanced TBI's. This work will build upon my initial research in the area [8, 9, 11, 14]. The research will provide an ideal training ground for a PhD student. He/she will learn about the design, conduct and analysis of experiments and the use of sound. There will also be great scope for individual research by an able student because there are many open questions in this area. The student must have skill in human-computer interaction and some knowledge of Macintosh programming. I currently have several undergraduate project students using sound and would hope to recruit one of these.

The project requires one Apple Macintosh computer. The machine will be used for the design and production of the sounds. The Macintosh is the standard environment used by music and sound professionals to design sounds and the best tools for the creation and manipulation of sounds are available on this platform. This machine will also be used for the production of papers, software and the running of experiments to test out new TBI's. Our industrial partners will provide test-bed equipment for us to build our test TBI's. The sounds we create on the Macintoshes will be used on this equipment.

The department of Computing Science already has an established Macintosh network and there are technical support staff to deal with any problems. However, there is no backup system in the department so I have requested this as part of the project. It will be very important to have backups of the sounds, code and papers from the project.

We also require a MIDI sound module plus sound-output hardware. Together, these will be used to generate the sounds and then present them to experimental participants for evaluation.

Travel funds are necessary to ensure the adequate dissemination of our work. As the work is based on sound it is very difficult to describe without demonstrations. Presenting the work at conferences is the most suitable way of demonstrating our results to academia and industry. I have also requested funding for a workshop at the end of the project so that our knowledge can be passed on to industry and academia.

PART 2: PREVIOUS RESEARCH AND TRACK RECORD

I have been a lecturer in the Department of Computing Science at the University of Glasgow since October, 1995. From September, 1994 until October 1995 I was an ERCIM Research Fellow working at VTT in Finland and SINTEF DELAB in Norway. Previous to that, I worked on my PhD at the University of York, completing it in August, 1994. In 1992 I was awarded the Gibbs-Plessey prize for postgraduate research in computer science.

I have been on the program committees of APCHI'96 and BCS HCI'95 and HCI'96 and was on the international committee for ACM CHI'95 and CHI'96. I was on the review committee for ICAD'94 and ICAD'96. I am also involved in designing the sound component for PREMO, the ISO standard for the presentation and realisation environment for multimedia objects. I am a member of the ACM, SIGCHI and the BCS HCI groups.

For my PhD research [6] I investigated the design and use of non-speech sounds in human-computer interfaces. I carried out a series of detailed experiments to discover the most effective ways of constructing earcons [14, 17]. From this work a set of guidelines for the construction of earcons was produced. I then used earcons to correct usability errors in standard graphical widgets (buttons and scrollbars) for sighted and partially sighted users. Results showed statistically significant improvements in time to recover from errors, time to complete tasks and reduced workload [12, 15]. Users also did not find the sounds annoying; in fact they significantly preferred the new widgets to the old ones. This work provides a solid foundation for the research proposed here.

I extended this research during my ERCIM Research Fellowship spent at two major European research centres, both of which are investigating the effective uses of sound. Whilst at VTT in Helsinki, Finland, I worked on using earcons as navigation aids in interfaces where graphical feedback is not available (for example, telephone-based interfaces and those for visually disabled people) as part of the TIDE ACCESS project [9]. The results of the work are discussed above. This is directly relevant to the project described here.

A second experiment investigated the use of sound to enhance single-switch scanning input [10]. This style of interaction is used by people with severe physical disabilities who cannot operate more complex devices like mice. Automatic scanning of items is used to select the required item. The layout of the scanned items is spatial but the selection task is temporal. My hypothesis was that sound could help as it is better suited for temporal tasks than graphical feedback. Sounds were designed to exploit the rhythmic nature of the scanning process to allow more rapid and accurate selection when scanning. Preliminary results were favourable. This work again improved general understanding of the use of earcons to improve human-computer interaction.

At SINTEF DELAB in Trondheim, Norway I worked in the acoustics department further developing earcons. I and colleagues investigated the use of non-speech sound to present constant, simultaneous feedback in human-machine interfaces. In some situations, drivers of commercial or military vehicles cannot easily look at dials and gauges whilst driving because they must concentrate on the view out of the windscreen. Sound was used to give information about constantly changing fuel, engine temperature and oil pressure levels in a simulation of this environment. The aim of the research was to investigate the problems of presenting many sound sources simultaneously and see if participants would perform better if they did not have to look at graphical gauges representing the same information. The results showed that workload was significantly reduced by the addition of sound; participants could concentrate on driving without having to look at the fuel, temperature and pressure gauges. However, this did not affect the participant's ability to respond to spoken messages and did not increase annoyance. The knowledge gained from this experiment on the design of earcons will feed into the work described in this proposal.

Local Expertise

There is much local expertise at Glasgow that can provide feedback and support for this proposal. I am currently part of the internationally-renowned GIST Group at Glasgow. This multidisciplinary group contains skills in all areas of human-computer interface design, from computing science to psychology. Members will be able to provide feedback on all areas of the research. For example, Dr Steve Draper, Department of Psychology, has much experience in the design and evaluation of interfaces. His skills will be beneficial when designing experiments to test the sonically-enhanced interfaces and evaluating the results. Professor Derek Mcaulay, Department of Computing Science, has much expertise in sound technology. His existing EPSRC MINIM and SHEFC NetMuse projects aim to investigate the provision of high quality interactive distributed music services. Professor Mcaulay will be able to provide valuable advice on the problems of presenting sounds via computers. Dr Bill Findlay, Department of Computing Science, is an expert in music synthesis and the extraction of musical structure from digital audio data. He will be able to provide input on the design and control of the sounds used in the project.

Beneficiaries

My work has improved the usability of computer interfaces for a wide range of users. My research has shown significant benefits for sighted users using graphical user interfaces because it allows them to work faster and recover from errors more quickly. It also has potential benefits for partially-sighted users who will be able to use the enhanced graphical interfaces more easily. My research has improved the use of computers by physically disabled people by improving interaction techniques.

I have also developed a new method by which navigation cues can be given in non-visual interfaces. This work has great potential benefits for TBI's and other interfaces where graphics cannot be used. This also has strong benefits for blind users who need to access hierarchies of information such as when using files systems.

As well as users, interface designers have also benefited from my research. My work has shown how sounds can be integrated into graphical user interfaces and I have also proposed guidelines to help designers create effective sounds. This allows them to use sound in their interfaces more easily, and so improve usability, without needing to be experts in sound design. They will therefore be able to utilise the experimentally-proven advantages offered by sounds. Projects with Nokia, IBM and Telecom Sciences show this is the case.

REFERENCES

1. Aboaba, A. *Using sound in telephone-based interfaces*. MSc IT Thesis, University of Glasgow, 1996.
2. Alty, J.L. Can we use music in human-computer interaction? In *Proceedings of HCI'95* (Huddersfield, UK) Cambridge University Press, 1995, pp. 409-423.
3. Arons, B. and Mynatt, E. The future of speech and audio in the interface. *SIGCHI Bulletin* 26, 4 (1994), 44-48.
4. Blattner, M., Papp, A. and Glinert, E. Sonic enhancements of two-dimensional graphic displays. In *Proceedings of ICAD'92* (Santa Fe Institute, Santa Fe) Addison-Wesley, 1992, pp. 447-470.
5. Blattner, M., Sumikawa, D. and Greenberg, R. Earcons and icons: Their structure and common design principles. *Human Computer Interaction* 4, 1 (1989), 11-44.
6. Brewster, S.A. *Providing a structured method for integrating non-speech audio into human-computer interfaces*. PhD Thesis, University of York, UK, 1994.
7. Brewster, S.A. Using Non-Speech Sound to Overcome Information Overload. *Accepted for publication in Displays*, *Special Issue on Multimedia Displays* (1996).
8. Brewster, S.A., Raty, V.-P. and Kortekangas, A. *Representing complex hierarchies with earcons*. European Research Consortium for Informatics and Mathematics., 1995, Technical report, ERCIM-05/95R037.
9. Brewster, S.A., Raty, V.-P. and Kortekangas, A. Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy. In *Proceedings of HCI'96* (Imperial College, London, UK) Springer, 1996, pp. 169-183.
10. Brewster, S.A., Raty, V.-P. and Kortekangas, A. Enhancing scanning input with non-speech sounds. In *Proceedings of Assets'96* (Vancouver, Canada) ACM Press, 1996, pp. 10-14.
11. Brewster, S.A., Raty, V.-P. and Kortekangas, A. *Using earcons to provide navigational cues in a complex menu hierarchy*. Department of Computing Science, University of Glasgow, 1996, Technical Report, TR-1996-24.
12. Brewster, S.A., Wright, P.C., Dix, A.J. and Edwards, A.D.N. The sonic enhancement of graphical buttons. In *Proceedings of Interact'95* (Lillehammer, Norway) Chapman & Hall, 1995, pp. 43-48.
13. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. A detailed investigation into the effectiveness of earcons. In *Proceedings of ICAD'92* (Santa Fe Institute, Santa Fe) Addison-Wesley, 1992, pp. 471-498.
14. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of INTERCHI'93* (Amsterdam, Holland) ACM Press, Addison-Wesley, 1993, pp. 222-227.
15. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. The design and evaluation of an auditory-enhanced scrollbar. In *Proceedings of CHI'94* (Boston, Ma.) ACM Press, Addison-Wesley, 1994, pp. 173-179.
16. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of HCI'95* (Huddersfield, UK), 1995, pp. 155-159.
17. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. Parallel earcons: Reducing the length of audio messages. *International Journal of Human-Computer Studies* 43, 2 (1995), 153-175.
18. Cohen, J. Monitoring background activities. In *Proceedings of ICAD'92* (Santa Fe Institute, Santa Fe) Addison-Wesley, 1992, pp. 499-532.
19. Edwards, A.D.N. Soundtrack: An auditory interface for blind users. *Human Computer Interaction* 4, 1 (1989), 45-66.
20. Edwards, A.D.N., Eds. *Extra-Ordinary Human-Computer Interaction*. Cambridge University Press, Cambridge, UK, 1995.
21. Edworthy, J., Loxley, S. and Dennis, I. Improving auditory warning design: Relationships between warning sound parameters and perceived urgency. *Human Factors* 33, 2 (1991), 205-231.
22. Gaver, W., Smith, R. and O'Shea, T. Effective sounds in complex systems: The ARKola simulation. In *Proceedings of CHI'91* (New Orleans) ACM Press, Addison-Wesley, 1991, pp. 85-90.
23. Goldstein, M. and Bretan, I. The AlphaWheel: Character input in minimal interfaces. In *Adjunct Proceedings of HCI'96* (Imperial College, London, UK) Springer, 1996, pp. 69-72.
24. Maguire, M. A human-factors study of telephone developments and convergence. *Contemporary Ergonomics* (1996), 446-451.
25. Patterson, R.D. *Guidelines for auditory warning systems on civil aircraft*. Civil Aviation Authority, London, 1982, CAA Paper, 82017.
26. Resnick, P. and Virzi, R. Skip and Scan: Cleaning up telephone interfaces. In *Proceedings of CHI'92* (Monterey, California) ACM Press, Addison-Wesley, 1992, pp. 419-426.
27. Rosson, M.B. Using synthetic speech for remote access to information. *Behaviour Research Methods, Instruments and Computers* 17, 2 (1985), 250-252.
28. Schumacher, R.M., Hardzinski, M.L. and Schwartz, A.L. Increasing the usability of interactive voice response systems. *Human Factors* 37, 2 (1995), 251-264.
29. Stevens, R.D., Brewster, S.A., Wright, P.C. and Edwards, A.D.N. Providing an audio glance at algebra for blind readers. In *Proceedings of ICAD'94* (Santa Fe Institute, Santa Fe) Santa Fe Institute, 1994, pp. 21-30.
30. Wolf, C., Koved, L. and Kunzinger, E. Ubiquitous Mail: Speech and graphical interfaces to an integrated voice/email mailbox. In *Proceedings of Interact'95* (Lillehammer, Norway) Chapman & Hall, 1995, pp. 247-252.
31. Yankelovich, N., Levow, G. and Marx, M. Designing SpeechActs: Issues in speech user interfaces. In *Proceedings of CHI'95* (Denver, Colorado) ACM Press, Addison-Wesley, 1995, pp. 369 - 376.

PART 3: DIAGRAMMATIC PROJECT PLAN

	Stephen Brewster	Research Student
Year 1	Leading information gathering exercise and liaising with industrial partners Supervise RS and help design the navigation experiments	Design and run set of experiments to investigate use of earcons to for navigation cues. Compare decimal representations to previous results First year report
	Write paper describing the results of the experiments	
Year 2	Support work of the RS, help with the design of the guidelines for the use of sound Evaluate the earcon guidelines with our industrial partners	Experimentally test navigation earcons in real-world TBI then develop guidelines for creating navigation cues Develop new interaction techniques using enhanced audio cues
	Write paper describing guidelines for adding sound cues	
Year 3	Work with industrial partner to design the final test-bed TBI the RS will enhance Support RS in writing up thesis	Evaluate navigation cues and new interaction techniques in enhanced TBI. Plan thesis Write up thesis
	Write papers on new interaction techniques and final evaluation	
	Work on final report and plan career development for RS	
	Submit Final Report	Submit Thesis