Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy

Stephen Brewster[1], Veli-Pekka Raty[2] & Atte Kortekangas[3]

[1]Department of Computing Science, The University of Glasgow, Glasgow, G12 8QQ, UK. Tel: +44 (0)141 330 4932, Email: stephen@dcs.gla.ac.uk, Web: http://www.dcs.gla.ac.uk/~stephen/

[2]VTT Information Technology, Kanslerinkatu 12 B, Tampere, FIN-33101, Finland. Tel: +358 31 316 3328, Email: veli-pekka.raty@vtt.fi

[3]VTT Information Technology, Tekniikantie 4B, Espoo, FIN-02044 VTT, Finland. Tel: +358 0 456 4311, Email: atte.kortekangas@vtt.fi

We describe an experiment to discover if structured audio messages, earcons, could provide navigational cues in a menu hierarchy. A hierarchy of 27 nodes and four levels was created with sounds for each node. Participants had to identify their location in the hierarchy by listening to an earcon. Results showed that participants could identify their location with over 80% accuracy, indicating that earcons are a powerful method of communicating hierarchy information. Participants were also tested to see if they could identify where previously unheard earcons would fit in the hierarchy. The results showed that they could do this with over 90% accuracy. These results show that earcons are a robust and extensible method of communicating hierarchy information in sound.

Keywords: Earcons, auditory interfaces, non-speech audio, navigation, menus, phone-based interaction.

1 Introduction

This paper describes an experiment to investigate the ability of non-speech audio to provide navigational cues in a hierarchical menu structure. In some situations graphical feedback cannot be used to provide these cues. In completely auditory interactions, such as telephone-based interfaces or those for visually disabled people, it is impossible to use graphical cues. In other systems where graphical feedback is available, the display may already be completely occupied by information that extra graphical cues would hide. For example, an interface for people with speaking difficulties who need to access a library of pictographic images. The link between these different situations is that a hierarchical structure must be

represented without graphics. In this paper we describe an experiment to test sounds to see if they could convey the necessary hierarchy information.

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, for example Blattner *et al.* (1992), Brewster (1994) and Gaver *et al.* (1991). Non-speech sound is an important means of communication in the everyday world and the benefits it offers should be taken advantage of at the interface. Such *multimodal* interfaces allow a greater and more natural communication between the computer and the user. They also allow the user to employ appropriate sensory modalities to solve a problem, rather than just using one modality (usually vision) to solve all problems.

Sound has many advantages. It is omni-directional so the user does not have to concentrate on a particular part of the display to perceive it. In fact, he/she does not even have to be looking at the display. Sound is also attention grabbing and can be effectively used to indicate problems to users. It can work alongside synthetic speech in purely auditory interfaces or be integrated with graphical feedback. Often in graphical interfaces more and more information is displayed on-screen. This can result in overload and important information may be missed. One way to overcome this problem is to use sound. Important information can be displayed on the screen and other information in sound, reducing overload of the visual sense. Brewster (1994) showed that by adding sound to a graphical interface both the time taken to complete certain tasks and the time taken to recover from errors could be reduced.

1.1 Telephone-based interfaces

One important reason for using non-speech sound is to represent menu structures in interfaces where visual feedback is not possible, for example telephone-based interfaces (phone banking) or interfaces for visually disabled people. In a telephone-based interface a user might call the bank and navigate through a hierarchy of voice menus to find the service required. One problem is that the user can get lost in the hierarchy. As Yankelovich *et al.* (1995) say: “These interfaces, however, are often characterized by a labyrinth of invisible and tedious hierarchies which result when menu options outnumber telephone keys or when choices overload users’ short-term memory”. The communication channel is very limited so little feedback can be given to the user about his/her current location. The more navigation information that is given, the more it obstructs the information the user is trying to access.

Rosson (1985) investigated such a hierarchical phone-based interface to give travel/visitor information. A user could call the system and move through the hierarchy to find information such as addresses and phone numbers of restaurants. She describes one common problem (p 251): “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 (1996) and Stevens *et al.* (1994) when designing navigation cues in a system to provide non-visual access to mathematics. 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 low quality of 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).

We suggest an alternative solution: Structured non-speech audio messages. These would provide a hierarchical system of sounds that could be used to represent a menu hierarchy. 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. This is a similar approach to that taken by Stevens (1996) and Stevens *et al.* (1994) in the Mathtalk system. This system displays algebra to blind mathematicians. Non-speech sounds give the listener information about their location in a mathematical structure. They do this without interfering with the synthetic voice presenting the mathematics. The cues are also much shorter than an equivalent voice message. Speech and non-speech sounds are different in the same way as text and graphics. If carefully designed, they can also be used together to provide complementary information.

1.2 Navigation in a communicator device

A second reason for investigating the use of sound to represent hierarchies is to present navigational cues in an interface for people speech impairments for the TIDE ACCESS Project 1001. The aim of this project is to create a mobile communication device for people who cannot speak. People will use the device to create messages they want to communicate and then say these messages via synthetic speech. These types of users often use pictographic languages, for example Bliss (Baumgart *et al., 1990*), to communicate. The pictures represent words or actions and can be combined to create complex messages.

An experienced user with a wide vocabulary may need access to a large number of symbols. It is impossible to display all of the pictures on the screen at the same time. One way around this problem is to use a hierarchy of symbols but this can lead to users getting lost, in the same way as described above.

Graphical feedback could be used to give extra navigational information, for example a map could show the current position in the hierarchy. However, this would take up valuable screen space needed for the pictographic symbols and would also require the user to look at the map when he/she really wanted to look at the symbols. An alternative would be to use different colours to show the different levels. Unfortunately, colour is already used to show other groupings within the symbols such as nouns/verbs/adjectives. We suggest the use of sound; it can be used to give information about one's location in the hierarchy without taking up screen-space. Synthetic speech cannot be used as it would conflict with the message being created. However, non-speech would not suffer from this problem.

Lack of screen space is not only a problem in interfaces for disabled people. Visual displays can only hold so much information. If an interface designer tries to display more and more

information then some of it will not fit on the screen without hiding information already there. Blattner *et al.* (1992) discuss this problem with computerised maps. Only so much information can be displayed before the underlying map is obscured. If additional information must be displayed on a map, space must be allocated for it and eventually a saturation point will be reached. Blattner *et al.* suggested that sound could be used to avoid these problems.

2 Earcons

The non-speech sounds used for this investigation are based around structured audio messages called *Earcons* (Blattner *et al.,* 1989 and Brewster, 1994). Earcons are abstract, synthetic tones that can be used in structured combinations to create sound messages to represent parts of an interface. Detailed investigations of earcons by Brewster *et al.* (1993) showed that they are an effective means of communicating information in sound.

Earcons are constructed from motives. These are short rhythmic sequences that can be combined in different ways. The simplest method of combination is concatenation to produce *compound earcons*. By using more complex manipulations of the parameters of sound (such as timbre, register, intensity, pitch and rhythm) *hierarchical earcons* can be created (Blattner *et al., 1989*) which allow the representation of hierarchical structures.

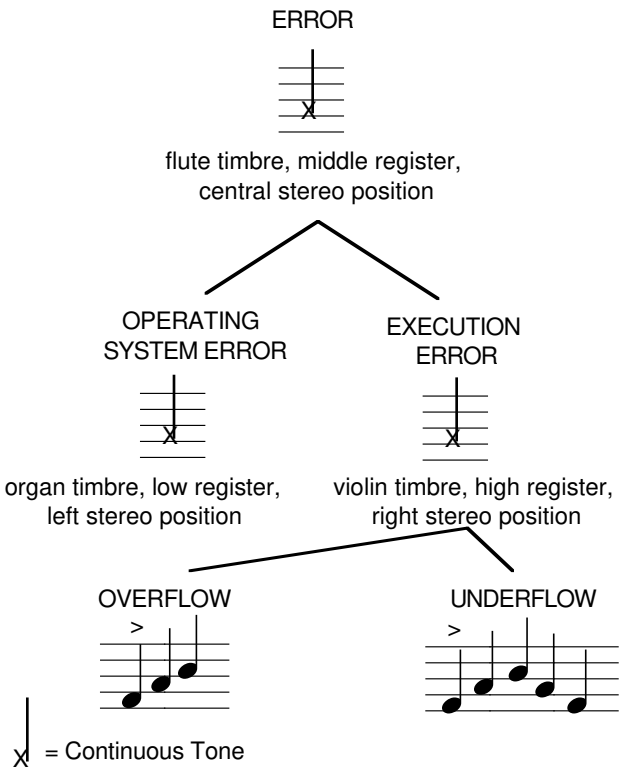

Figure 1: A hierarchy of earcons representing errors (adapted from Blattner *et al.*, 1989).

Figure 1 shows a simple hierarchy of earcons using these types of manipulations. Each earcon is a node on a tree and inherits all the properties of the earcons 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 fictitious family of errors. It has a flute timbre (a 'colourless' instrument), a middle register and a

central stereo position. The structure of the earcon from level one is inherited by level two and then changed. At level two the sound is still continuous but non-neutral timbres are used (in the figure organ and violin). Register is changed so that it matches a conventional musical layout (low register on the left, high on the right) and stereo position reflects the layout of the hierarchy, for example the node on the left has a left stereo position. 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, register and stereo position from the level above. Other levels can be created by using parameters such as tempo or effects.

Using earcons, this hierarchy is easily extensible. For example, to add another major category of errors all that is needed is a new timbre. This could then be given a middle register and a middle stereo position and be placed between the operating system and execution errors in the diagram. 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 creating hierarchies.

An alternative method of presenting information in sound is *Auditory Icons*, for example Gaver (1989) and Gaver (1993). These are based around sounds sampled from the everyday environment (such as metallic or wooden sounds) rather than the musical sounds of earcons. Unlike abstract earcons, these sounds have an intuitive link to the object or action being represented. For the creation of hierarchies earcons have the advantage that the sounds can be varied in a systematic way (as shown in the description above). Auditory icons are based on natural sounds and so cannot be as easily manipulated (if an auditory icon is manipulated too much it may no longer sound like its natural equivalent). Gaver is working on this problem (Gaver, 1993) to allow the controlled variation of auditory icons. However, at the present time it is not possible.

2.1 Previous attempts to use earcons to present hierarchy information

Barfield *et al.* (1991) carried out experiments where they used earcons to aid navigation through a menu hierarchy. The earcons they used were very simple, just decreasing in pitch as a participant moved down the hierarchy. The sounds lasted half a second. They describe them thus (p 104): "…the tones were played with a harpsichord sound and started in the fifth octave of E corresponding to the main or top level of the menu and descended through B of the fourth octave". These sounds did not fully exploit all the advantages offered by earcons (for example, they used neither rhythm nor timbre differences) and did not improve user performance in the experimental task.

Brewster (1994) and Brewster *et al.* (1993) also tested the ability of earcons to present hierarchical information in sound. In two detailed experiments they showed that, with careful design of the earcons, hierarchy information could be presented effectively. They used earcons to represent a small family of files, folders and applications and manipulated timbre, register, rhythm and pitch to create their hierarchy. Their results showed that 80% recall rates could be achieved for hierarchical earcons even with non-optimal training. This work indicated that earcons could be used to represent hierarchies. However, the hierarchy used was very simple with only nine nodes.

3 The experiment

The aim of the experiment described here was to discover if a larger hierarchy could be represented by earcons. Figure 2 shows the hierarchy used. It had 25 nodes on four levels with four missing nodes on Level four (two of which are marked as A and B). This made a

hierarchy of 27 nodes, three times larger than that tried previously. It was loosely based on the structure of the file system on the first author's computer.

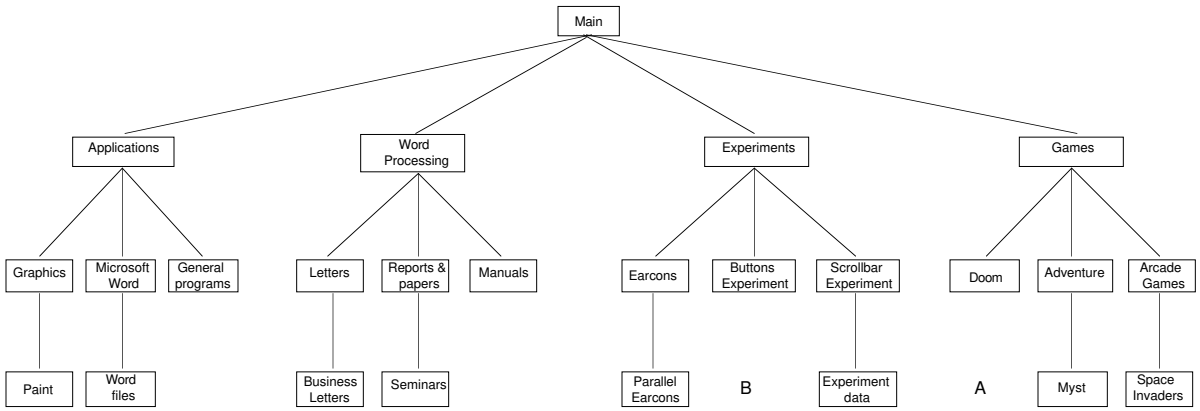

Figure 2: The file-system hierarchy used in the experiment. A and B show the two new earcons presented to participants during testing.

3.1 Hypotheses

The main hypothesis was that participants should be able to recall the position of a node in the hierarchy by the information contained in an earcon. If this was correct then high overall rates of recall would be expected.

Participants should also be able to listen to an earcon and position it in the hierarchy even if they have not heard it before by using the rules from which the earcons were constructed. This would be demonstrated by high rates of recognition when participants were presented with new earcons.

3.2 Participants

Twelve volunteer participants were used. They were a mixture of computer science students at Helsinki University of Technology and members of staff at VTT Information Technology. All were familiar with computers and computer file systems.

3.3 Sounds used

The earcons were designed using the guidelines proposed by Brewster *et al.* (1995). The sounds were all played by HyperCard on an Apple Macintosh via MIDI using a Yamaha TG100 sound synthesiser and presented to participants via loudspeakers. The sounds used at each level of the hierarchy will now be described:

Level 1: For the top level of the hierarchy ('Main' in Figure 2) a constant sound with a flute timbre was used (see Table 1). It had a central spatial location and a pitch of D3 (261Hz). A flute timbre was used at it is a pure sound close to a 'timbreless' sinewave. The earcon was designed to be neutral sounding.

Level 2: Each family had a separate timbre, register and spatial location. Table 1 shows these. Register was lowest on the left and highest on the right following the conventional

musical pattern. The stereo position of the earcons also moved from left to right mirroring their position in the hierarchy.

Nodes	Timbre	Stereo position	Register
Main	Flute	Centre	D_3
Applications	Electric organ	Far left	C_4
Word Processing	Violin	Centre left	C_3
Experiments	Drum/synthesiser	Centre right	C_2
Games	Trumpet	Far right	C_1

Table 1: The timbre, spatial location and register for Levels 1 and 2 of the hierarchy.

The continuous sound was inherited from the Level 1 earcon but the instrument, pitch and stereo position were changed. Three parameters were used so that if the listener could not remember which instrument went with which node he/she could still use register or stereo position.

Stereo position was used in the earcons even though it is not available in telephone-based interfaces. We did this because it is be available in interfaces for disabled people based on personal computers, such as the portable communicator device described above or in interfaces for the blind. For these two situations we wanted to use all of the techniques available to maximise usability.

Level 3: At this level rhythm was used to differentiate the nodes. Each left node had one rhythm, each centre node another rhythm and each right node another. Figure 3 shows the rhythms used. From Figure 2 'Graphics', 'Letters', 'Earcons' and 'Doom' all had the left node rhythm, 'Microsoft Word', 'Reports & Papers', 'Buttons Experiment' and 'Adventure' were centre nodes and 'General Programs', 'Manuals', 'Scrollbar Experiment' and 'Arcade Games' were right nodes. Each of these rhythmic groups repeated continuously once every 2.5 seconds. As Figure 3 shows, the first note in each group was accented. The last note of each group was also lengthened slightly. These two help make each group into a complete rhythmic unit (Brewster *et al.*, 1995).

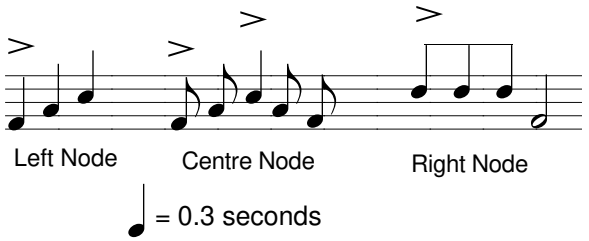

Figure 3: The rhythms used for Levels 3 and 4 of the hierarchy.

At this level the earcons inherited timbre, spatial location and register from Level 2. This meant, for example, that 'Graphics' used the left node rhythm described in Figure 3 and it was played with an electric organ timbre, on the left side of the stereo space and in the register of C_4. 'Letters' used the same rhythm but, in this case, the timbre was a violin, stereo position was centre left and the register was C_3.

Level 4: A faster tempo was used to differentiate the items. The rhythmic units from Figure 3 now repeated once every second. In addition to this the effects reverb and chorus were applied to all of the earcons. These gave the earcons a much fuller sound. This time rhythm

was inherited from Level 3. Each of the nodes in Level 4 used the same rhythm as its parent node but the earcons were repeated more frequently.

3.4 Experimental design and procedure

As shown in Figure 2, the hierarchy was based on a computer file system. This was an experiment to test the use of earcons to represent a hierarchy and a file system was a convenient hierarchy to use, this paper does not suggest that each directory and sub-directory in a real file system should have a sound.

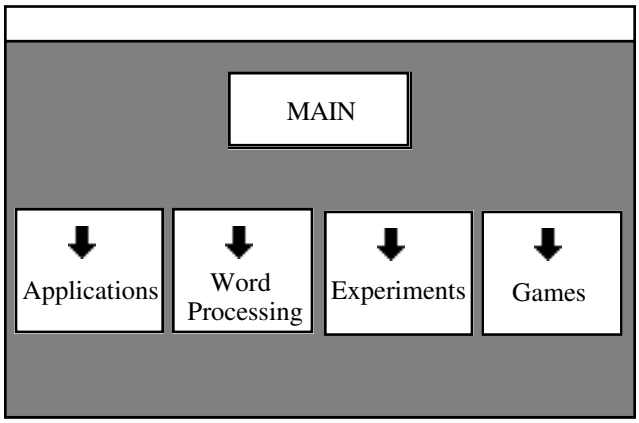

Figure 4: The top level of the hierarchy. The arrows show the direction of movement possible.

The hierarchy was constructed in a HyperCard stack. Figure 4 shows the screen of the top level of the hierarchy ('Main'). Each of the boxes in Figure 2 was a card in the stack. Buttons were provided for going up and down levels in the hierarchy and also for going left and right across the same level. As soon as a card was selected its sound started to play and continued until another card was selected.

3.5 Training

The training was in two parts. In the first part the experimenter showed the participant each of nodes of the hierarchy in turn and played the associated earcon. This was done once only. The structure of the earcons at each level was fully explained. In the second part of the training participants were given five minutes to learn the earcons by themselves with no help from the experimenter. The training was short and simple allowing us to see if the rules by which the earcons created were obvious. Users of a communicator device of the type described above might only receive very simple training and have to learn the rest of the system by themselves. Users of a telephone-based system might get a short amount of training when they signed-up for a new telephone service and again have to learn it by themselves. The simple training used here would allow us to test these situations.

During the training participants could look at a map of the hierarchy (similar to Figure 2). The aim of the experiment was not to test the participants' abilities to learn hierarchies but to test their ability to learn the earcons. Instructions were read from a prepared script.

3.6 Testing

The participants heard fourteen earcons during testing. These were randomly selected from all of the sounds in the hierarchy. The same set of earcons was presented to each of the participants. Twelve of the sounds were ones that participants had heard during the training. The last two earcons presented were new ones (marked A and B in Figure 2). These were earcons for gaps in the hierarchy and were constructed using the same rules as the other earcons. Table 2 shows which node represented which question. An earcon was played and the participants then had to choose where the it fitted into the hierarchy. The hierarchy was represented on screen. Again, the aim of the experiment was to test the participants' knowledge of the earcons, not their ability to learn hierarchies. None of the names of the nodes were included in the picture of the hierarchy to avoid any help they might have provided.

Question	Level	Node
Q1	2	Word processing
Q2	4	Space invaders
Q3	2	Experiments
Q4	4	Paint
Q5	2	Games
Q6	3	Doom
Q7	4	Word files
Q8	4	Parallel earcons
Q9	3	Graphics
Q10	3	Reports & papers
Q11	4	Business letters
Q12	3	Microsoft word
Q13	4	A
Q14	4	B

Table 2: The node and level in hierarchy for each of the questions. This is the order that the questions were presented to participants.

4 Results

4.1 Overall

The overall recall rate of the earcons was high: 81.5% of the earcons were correctly recalled. Figure 5 shows the percentage of correct answers for each question. An analysis was undertaken to find out if any of the earcons were less well recalled than any others. A one-factor ANOVA was performed on the scores per question and it showed an effect ($F(13,154)=2.13$, $p=0.01$). In order to find out where the effect was (and so which were the worst-recalled earcons) *post hoc* Tukey HSD tests comparing each question were performed (see Table 3). The three worst recalled earcons were from questions 2, 4 and 11. They were all from Level 4 of the hierarchy. The nodes were: 'Space Invaders', 'Paint' and 'Business Letters'. However, the other Level 4 nodes 'Word Files' and 'Parallel Earcons' were significantly better recalled than these. 'Paint' was recalled worst of all.

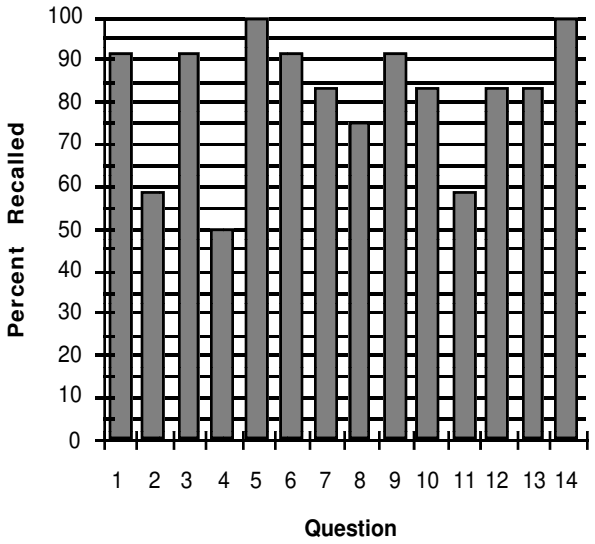

Figure 5: Recall rates for each of the 14 questions.

4.2 Recall of components

Each of the participants' answers were broken down to find where the problems occurred. There were three mistakes that could be made. Participants could mistake the *family* of an earcon: For example, whether it was from 'Applications', 'Word Processing', 'Experiments' or 'Games'. They could also mistake the *node* an earcon referred to: Whether it was a left node, a centre node or a right node. Finally, participants could mistake the *level* of an earcon: Whether it was from Level 1, 2, 3 or 4. From the overall data obtained the scores were broken into three. If participants got a question completely right they received three marks, if they got two parts right then two marks were awarded, etc. From this analysis it was possible to see where mistakes occurred.

Questions	Question 4	Questions 2 & 11
q1	Q(154)=5, p=0.01	Q(154)=4, p=0.01
q2	Q(154)=1, p=0.01	--
q3	Q(154)=5, p=0.01	Q(154)=4, p=0.01
q4	--	--
q5	Q(154)=6, p=0.01	Q(154)=5, p=0.01
q6	Q(154)=5, p=0.01	Q(154)=4, p=0.01
q7	Q(154)=4, p=0.01	Q(154)=3, p=0.01
q8	Q(154)=3, p=0.01	Q(154)=2, p=0.01
q9	Q(154)=5, p=0.01	Q(154)=4, p=0.01
q10	Q(154)=4, p=0.01	Q(154)=3, p=0.01
q11	Q(154)=1, p=0.01	--
q12	Q(154)=4, p=0.01	Q(154)=3, p=0.01
q13	Q(154)=4, p=0.01	Q(154)=3, p=0.01
q14	Q(154)=6, p=0.01	Q(154)=5, p=0.01

Table 3: Tukey HSD results comparing the three worst questions. Question 4 was significantly worse than any other question. Questions 2 & 11 were worse than all except question 4.

There were no significant differences between the recall rates of family, level and node. Participants recalled 80% of the family and level components and 76% of the node components. A one-factor ANOVA on the family, level and node scores showed no significant difference between these ($F(2,33)=1.27$, $p=0.291$).

4.3 New earcons

As mentioned above, two new, previously unheard, earcons were presented to the participants during testing. These are marked by the letters A and B in Figure 2. They were both from Level 4 and were constructed using the same rules as the rest of the earcons. By using the rules participants should have been able to identify where in the hierarchy the sounds belonged.

The new earcons were very well recognised (see questions 13 and 14 in Figure 5): 91.5% were recognised correctly. Ten out of twelve participants recognised the earcon for A in Figure 2 and all the participants recognised the earcon for B. This indicated that the participants were able to use the rules to work out where an unheard earcon belonged.

5 Discussion

5.1 Overall

The overall recall rate of the earcons was good. The participants recalled 81.5% of the earcons after being trained for only a short time. This demonstrates that a listener could use earcons as navigation aids in a menu hierarchy. This was also possible after only a short amount of training. Participants had the sounds described to them once only and then had five minutes to listen to the sounds. This shows that with careful design, earcons can easily be learned. One of the advantages claimed for auditory icons (described above) is that they are easy to learn because they use natural, everyday sounds. The research described here, combined with the earlier work of Brewster (1994) and Brewster *et al.* (1993), suggests that there may be no significant difference between auditory icons and earcons in this respect.

The results obtained here are approximately the same as the rate of recall in phase I of the second earcon experiment described by Brewster (1994). As mentioned above, the hierarchy used in that experiment was much simpler but the training and testing were less efficient (for example, the earcons were presented in a random order during training). In the experiment described here the training was more structured but the hierarchy was three times larger, leading to the similar recall rates.

The three worst recalled earcons were from Level 4 of the hierarchy. However, the other two earcons from Level 4 were significantly better recalled than these. A detailed analysis of the results gave no clear indication of what might have caused the difficulty. The problem may have been that these earcons were at the bottom of the hierarchy so participants had to remember all of the earcon construction rules to work out their location. This left more opportunity for mistakes than with the recall of earcons higher-up in the hierarchy where less had to be remembered. The manipulations chosen for Level 4 (tempo, chorus and echo) could be further exaggerated to more clearly indicate the differences between items. Further investigation is necessary.

5.2 New earcons

The ability of the participants to identify the location of previously unheard earcons was very good (91.5%). They were able to use the rules for constructing the earcons to work out where a new earcon belonged. This showed that the listeners had learned the rules well. As discussed above, earcons provide an easily extensible system for adding sound for hierarchies. This research has shown that, if new nodes were added to an existing hierarchy then listeners would not need to be retrained on the sounds, they could use their existing knowledge of the rules to understand the new earcons. This is a very promising result for the use of earcons as hierarchical navigation aids.

The new sounds were from one of each of the main families, or major sub-trees, of the hierarchy so the participants could use all of the earcon creation rules to identify them. There were also only four missing ones so that their choice was reduced. For example, if they could only remember the timbre for the sub-tree they would still be able to identify the new earcon because each new earcon had a different timbre. This test could have been made more difficult by having two missing earcons from the same sub-tree. This would have made it harder to identify the missing one as there would be fewer differences between them. However, the result does show that participants were able to use the rules to locate the new earcons. This indicates that they are a very robust method of communicating hierarchy information in sound.

6 Future work

These earcons will now be used for the TIDE ACCESS Project (they will be extended where necessary using the techniques described above). The sounds will be played quietly, in the background, when the user is in the hierarchy of pictographic symbols. The sounds will change when he/she moves up, down, left or right to a new node, indicating the new position in the hierarchy. The change in the earcon (plus a small increase in intensity) will capture the listener's attention and indicate the new location. The earcon will then fade back to a lower intensity and recede into the background of the listener's consciousness. The user will only attend to the sound again if he/she gets lost and wants navigation information.

For telephone-based interfaces this work must be extended. In a telephone-based interface the quality of the sounds would be reduced due to the limited bandwidth of telephone equipment. There would also be no stereo effects available. The next stage of the research will be to conduct the experiment again but present the sounds as if they were coming from a telephone. The results would then show what recall would be like under these more difficult conditions. The earcons could then be re-designed if necessary to maximise recall.

7 Conclusions

The aim of the experiment described here was to discover if earcons could provide navigational cues in a system of hierarchical menus. The results have shown that a 27 node hierarchy can easily be represented using earcons. A recall rate of 81.5% was achieved after only a short period of training. This indicates that they are a good way of providing navigational information for hierarchies. Listeners can hear an earcon and tell where it comes from in the hierarchy and hence where their location. Listeners could also recognise new earcons that had not been heard before (with 91.5% accuracy) by using the rules from which the other earcons were constructed. This shows that users could easily learn those rules. Therefore, earcons are a robust and extensible method of communication. This

research has shown that earcons are an effective way of providing hierarchy information in interfaces where graphical information cannot easily be used.

Acknowledgements

This research was supported by ERCIM Research Fellowship 94-04. It was undertaken whilst the first author was at VTT Information Technology in Finland. Thanks to Peter Wright at the University of York, UK for helping with the statistical analysis.

References

Barfield, W., Rosenberg, C. & Levasseur, G. (1991). The use of icons, earcons and commands in the design of an online hierarchical menu. IEEE Transactions on Professional Communication, **34**, 101-108.

Baumgart, D., Johnson, J. & Helmstetter, E. (1990). *Augmentative and Alternative Communication Systems for Persons with Moderate and Severe Disabilities*. Baltimore, Maryland: Paul Brookes Publishing Co.

Blattner, M., Papp, A. & Glinert, E. (1992). Sonic enhancements of two-dimensional graphic displays. In Kramer (Ed.), *Auditory Display, sonification, audification and auditory interfaces. The Proceedings of the First International Conference on Auditory Display*, pp. 447-470, Santa Fe Institute, Santa Fe: Addison-Wesley.

Blattner, M., Sumikawa, D. & Greenberg, R. (1989). Earcons and icons: Their structure and common design principles. *Human Computer Interaction*, **4**, 11-44.

Brewster, S.A. (1994) *Providing a structured method for integrating non-speech audio into human-computer interfaces*. PhD Thesis, University of York, UK.

Brewster, S.A., Wright, P.C. & Edwards, A.D.N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In Ashlund, Mullet, Henderson, Hollnagel & White (Eds.), *Proceedings of INTERCHI'93*, pp. 222-227, Amsterdam: ACM Press, Addison-Wesley.

Brewster, S.A., Wright, P.C. & Edwards, A.D.N. (1995). Experimentally derived guidelines for the creation of earcons. In Kirby, Dix & Finlay (Eds.) *Adjunct Proceedings of HCI'95*, pp. 155-159, Huddersfield, UK.

Gaver, W. (1989). The SonicFinder: An interface that uses auditory icons. *Human Computer Interaction*, **4**, 67-94.

Gaver, W. (1993). Synthesizing auditory icons. In Ashlund, Mullet, Henderson, Hollnagel & White (Eds.), *Proceedings of INTERCHI'93*, pp. 228-235, Amsterdam: ACM Press, Addison-Wesley.

Gaver, W., Smith, R. & O'Shea, T. (1991). Effective sounds in complex systems: The ARKola simulation. In Robertson, Olson & Olson (Eds.), *Proceedings of CHI'91*, pp. 85-90, New Orleans: ACM Press, Addison-Wesley.

Rosson, M.B. (1985). Using synthetic speech for remote access to information. *Behaviour Research Methods, Instruments and Computers*, **17**, 250-252.

Stevens, R. (1996) *Principles for the design of auditory interfaces to present complex information to blind computer users*. PhD Thesis, University of York, UK.

Stevens, R.D., Brewster, S.A., Wright, P.C. & Edwards, A.D.N. (1994). Providing an audio glance at algebra for blind readers. In Kramer & Smith (Eds.), *Proceedings of ICAD'94*, pp. 21-30, Santa Fe Institute, Santa Fe: Addison-Wesley.

Yankelovich, N., Levow, G. & Marx, M. (1995). Designing SpeechActs: Issues in speech user interfaces. In Katz, Mack & Marks (Eds.), *Proceedings of CHI'95*, pp. 369 - 376, Denver, Colorado: ACM Press, Addison-Wesley.