An Investigation of Using Music to Provide Navigation Cues

Grégory Leplâtre and Stephen A. Brewster

Department of Computing Science, University of Glasgow
Glasgow, UK
Tel: (+44) 0141 339 8855
Fax: (+44) 0141 330 4913
Email: {gregory,stephen}@dcs.gla.ac.uk
Web: http://www.dcs.gla.ac.uk/~{gregory,stephen}

Abstract

This paper describes an experiment that investigates new principles for representing hierarchical menus such as telephone-based interface menus, with non-speech audio. A hierarchy of 25 nodes with a sound for each node was used. The sounds were designed to test the efficiency of using specific features of a musical language to provide navigation cues. Participants (half musicians and half non-musicians) were asked to identify the position of the sounds in the hierarchy. The overall recall rate of 86% suggests that syntactic features of a musical language of representation can be used as meaningful navigation cues. More generally, these results show that the specific meaning of musical motives can be used to provide ways to navigate in a hierarchical structure such as telephone-based interfaces menus.

1 Introduction

This paper describes an experiment to investigate the effectiveness of non-speech audio for conveying navigational cues in hierarchical structures such as telephone-based menus. Earcons (structured non-speech sounds [4]) have already been shown to be a powerful means to present a hierarchy in sound [7, 10]. Nevertheless, the need for representing bigger hierarchies has required more research. Musical sounds have a particular status among non-speech sounds. The meaning of a musical piece does not only belong to the way the piece is structured, it also and mainly belongs to the musicality perceived by individuals. We show that the design principles of earcons can be extended in reference to this property of musical sounds. Besides, this richness of musical semantics can be exploited using purely syntactic differences between sounds. A 25 node hierarchy of sounds was created without using differences between instruments, which have been proved to be the most efficient means of creating hierarchical earcons by Brewster *et al.* [10]. The background to this research is the improvement of interaction in telephone-based interfaces. In the following section, we address the issue of using non-speech sounds in this type of interface.

1.1 Telephone-Based Interfaces

The telephone is an ubiquitous device and is many people's primary method of entering into information infrastructures. According to Yankelovich *et al.*, [27], one major problem with this interaction is related to navigation [25]. As the multiple functions of telephone-based systems are accessible by means of a large complex menu, users have difficulties to reach the information they need [21, 26]. To facilitate users in their navigation tasks, an alternative to speech recognition lies in the use of non-speech sounds [9].

First of all, the telephone device itself allows only a limited form of interaction (small keypad, no screen, or a very small one). Schumacher *et al.* [23] suggest that the narrow channel of interaction consequent from the structure of this interface reduces the usability of telephone-based systems. Apart from the interaction problem caused by the physical structure of the interface, other interaction problems are related to the system features. Maguire [17] asserts that the multiplicity of available services and widened functionality of telephone systems are such that interaction with telephone-based devices needs to be addressed. Helmreich [14] claims that user interfaces are poor, so they do not allow users to access the wide range of functions of their telephone. Apart from the traditional use of telephone systems, telephone-based interfaces can also be a good medium of interaction with computers. Schmandt [22] points out that telephones indeed can be used as computer terminals. Zajicek and Powell [29] assert that using the telephone as a computer interface is a promising method to allow visually impaired people to use the web. Still, interaction needs to be improved since orientation and navigation remain sources of problems.

Roberts *et al.* [20] suggest that 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 interface itself. In the previous paragraph, we have highlighted the constitutive features that make telephone-based interfaces hard to use. Roberts *et al.* [20] propose that two means can broaden the narrow communication channel inherent to this interface:

- "One is to keep the same hardware but to incorporate much more intelligence in the system for both input and output: for example the system could recognize speech for input and synthesize speech for output, allowing the richness of language to increase the bandwidth".
- "The second method is to add to the modality of the bandwidth, allowing the user a visual channel by providing a display, and allowing richer input with pointing devices and other methods for inputting instructions".

In accordance with the second point, we can argue that one can "add to the modality" by exploiting the audio modality more extensively. Besides, this audio modality is certainly the dominant modality in the interaction with telephone devices. Now in telephony current standard guidelines do not provide the designer with much information about the use of non-speech sounds, apart from basic tones or combination of tones. Let us briefly describe the possibilities of using non-speech sounds in telephone environments.

There are two main paradigms for designing telephone-based interfaces: menu-based and conversational. In a graphical interface, the two main ways to process operations are executing commands from a prompt, or accessing the command navigating through a menu. Similarly, these paradigms are used in non-graphical telephone-based environments. The equivalent of command-based prompt systems is voice recognition based systems. In this case, there is no concept of navigation involved. The interaction is operated in a conversational fashion. Various authors have addressed this issue so far [18, 27, 28]. Graphical menus also find their equivalent in auditory menus.

By analogy with graphical display, both menu and conversational paradigms have strong and weak points in auditory display. Roughly, command-based systems requires more initiative from the user to complete a task, but expert users might prefer the flexibility of command-based systems though. Nonetheless, there is an interesting point that makes this analogy inaccurate. In auditory menus, the information is not parsed in the way it is in graphical menus. If the items of a graphical menu can be parsed quickly, items of an auditory menu have to be displayed sequentially in speech, which can take time. Williams *et al.* [25] assert that the length of menu prompts is a significant drawback in the use of this technology. The use of non-speech sound can contribute to reducing the time users take to listen to the sequences of menu items by providing relevant navigation cues.

1.2 Using Sound to Provide Navigational Cues

Earcons are structured non-speech sounds that can be combined, transformed, can inherit other earcons properties, and constitute an auditory language of representation (see [5] for an introduction to earcons, and [8] for more up to date information). Earcons are constructed from elementary objects of the language that will be referred to *motives*. By modifying the psychophysical parameters of the sounds from which motives are constructed, it is possible to create hierarchical earcons. Examples of hierarchical earcons have been developed by Blattner *et al.* [5]. For example, they created a message for an error class defined by a rhythmic pattern of unpitched sounds (e.g. clicks). Then different types of error could be represented by a melody matching this rhythm. In addition, Brewster *et al.* [7, 10] conducted different experiments that investigated the use of earcons to represent hierarchical structures. These experiments have shown that people could recall a 25-node hierarchy of earcons with good accuracy.

Earcons can also be used for other applications, for instance creating an auditory map [6]. In this last example, earcons were added to an interactive, digital aerial view of Lawrence Livermore National Laboratory. Sounds played as the cursor was moved over various buildings on the map, indicating access privileges, administrative units and computers. Because they are structured abstract sounds, earcons can be used for a wide range of applications. Gaver reviews most of these applications as well as other strategies used in auditory display in a recent report [13].

2 Design of the Sounds

2.1 Investigating New Principles for the Design of Hierarchical Earcons

The parameters on which the construction of hierarchical earcons is founded are based on psychoacoustical research [11]. Therefore, the most meaningful of them are psychophysical features of the sounds such as timbre, register, pitch or intensity, and no more strong low-level parameters can be expected to arise. So we have undertaken a different approach to find new valuable elements for the design of hierarchical sets of sounds. So far the information conveyed by earcons is related to the low-level parameters cited above. Now we can imagine that higher-level features of earcons could also convey pieces of information as *leitmotifs* do in music (using musical features in interfaces has been discussed in [1, 2]) Practically, our approach can be described from a linguistic viewpoint. Assuming that music is a language, it is constituted of a grammar and a vocabulary. Earcons currently mainly exploit the elements of the vocabulary of this language, using different instruments or timbres to convey information. But apart from rhythm, the syntactic richness offered by music has rarely been investigated. In this study, we chose to focus only on syntax, to see if we could take advantage of it to provide subjects with navigational cues.

2.1.1 Why Use 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 [1, 3, 4]. Blattner and Greenberg [4] 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”.

Smoliar [23] 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. 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 [1] 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”.

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” [4]. 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 [4] suggest as well that non-speech sounds like earcons play the role of chorus in Japanese Noh drama (in Noh drama, a chorus is part of a coded language that transmit information about the context of the dramatic situation), though we 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 [13] outlines in a recent study that auditory interfaces have so far drawn very little on the possibilities suggested by music. 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 relies in that the richness of music belongs to the richness of its meaning rather than in the richness of its structures.

2.1.2 Musicality, or The Essence of a Musical Language of Representation

“Perceptual psychologists assume that music is made of notes”. This assertion by Cook [12], talking about the way psychologists investigate our understanding of music, could well apply to the way music is used in auditory display. Indeed Lodha *et al.* [16] argue that as far as mapping data to a melody is concerned, “melody, in fact, is nothing but a succession of pitches”. Notes and their low-level parameters provide designers with practical means to control their auditory display. On the other hand leitmotifs [2], that can be either a simple melody or a harmonic progression, are examples of high level objects that can carry information. Alty and Vickers [2] have designed leitmotifs hierarchically to help users understand the structure of Pascal programs. A given melody can have different musical meanings, depending on the way it carries information [15]. Let us take the example of a simple melody constituted of four notes. If each of these notes carries information through the ways their timbre are modified, the musicality of the whole melody will be affected by the fact that users have to perform an analytical listening of each of the four notes instead of the melody itself. Here no advantage is taken of the fact that succession of notes was in fact a musical object, with a specific meaning, which is its musicality. On the contrary, if the melody is processed in the display as a whole, by modifying its rhythm for example, it will keep its fundamental musical meaning. Of course, it seems trivial in this basic example that melodies should not be decomposed as described. However, in the case of a more complex piece of music, it is not trivial to decide which decomposition is right and which one may damage the overall musicality of the piece. In other words, it is not trivial how a complex piece must be structured in order to convey information without destroying its musical unity. According to a previous study on the question [15], we believe that it is a matter to pay a great attention to the balance of non-informational elements of the piece as well as the informational ones. For instance, concerning the previous four notes melody, an element of the piece that should remain untouched is certainly the relative differences of note pitches that carry the musical meaning of the piece.

Interestingly, using musical meanings of sounds to convey information requires the user to categorise these sounds in his/her own way. The designer cannot help this categorisation. This is where individual differences are likely to appear, on the basis of individual capability to categorise music. The sounds involved in the experiment described in this paper have been designed to answer the following question: Can we represent a hierarchy in sound by using purely syntactic operations on abstract musical motives?

2.2 The Hierarchical Set of Sounds

To investigate the use of strictly syntactic combinations to convey information, we have constructed a hierarchical set of sounds using a single instrument: The piano. The 25 node hierarchy was the same as the one used by Brewster *et al.* [7, 10] to allow comparison of the results:

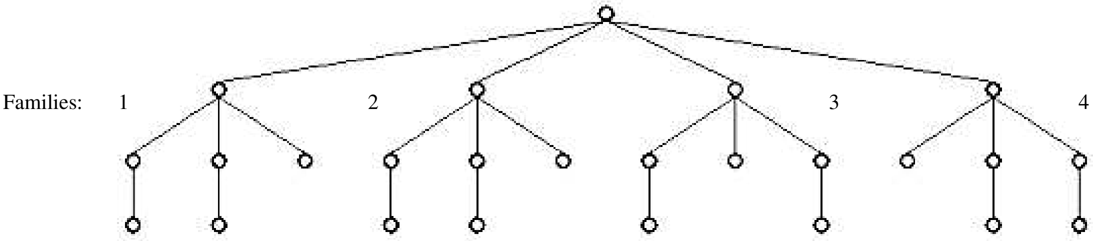

Figure 1: Hierarchical Structure Investigated[*].

The hierarchy represented in Figure 1 is composed of four main sub-trees that will be called families. Each of these families contain 6 nodes on 3 levels. The root of each family (level 1) is represented by a basic motive. The cues for the second level are represented by the level 1 motive 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 2 notes/chords, and 3 notes/chords for the third one. For each node of level 4 in the two first families, the cue is a chord succeeding the sound of the node above it. For each node of level 4 in families 3 and 4, the cue is a high register melody played in parallel with the sound of the node just above. In Table 1, the operator "+" symbolises the concatenation of two sounds, and "/" means that the sounds are played simultaneously. According to this notation, Table 1 (last column) shows that families 1 and 2 are structured sequentially, as opposed to families 3 and 4 whose third level sounds are played in parallel with the corresponding level 2 sounds. We intentionally divided the sounds in two these two categories (sequential and parallel) to know if individuals would experience difficulties to process mentally non-trivial piano streams.

Table 1 gives a rough idea of hierarchical system structure. The accurate structure of each sound appears more explicitly in Figures 2, 3, 4, and 5. The sounds are all available on the web at the following URL: www.dcs.gla.ac.uk/~gregory/web/experiment1. In the electronic version of this paper, these sounds can be reached via four links to each of the four families. We would like to warn the reader that the sounds described in this section have been designed for an experimental purpose only, which is investigated the efficiency of certain musical features to represent hierarchies in sound. Therefore, these sounds would not fit any sort of application as they are

Family \ Level	1	2	3
1	Low register C minor chord	1. + 1 medium register note 2. + 2 increasing medium register chords 3. + 3 increasing medium register chords	+ 1 high register chord
2	Medium register extract of a scale by tone	1. + 1 high register note 2. + 2 medium register decreasing chords 3. + 3 medium register decreasing chords	+ 1 low register chord
3	Short rhythmed tune	1. + 1 high register note 2. + 2 high register chords 3. + 3 high register chords	/ 1 higher register melody
4	Tune based on a II V I harmonic progression	1. + 1 medium register note 2. + 2 medium register increasing notes 3. + 3 medium register increasing notes	/ 1 higher register melody

Table 1: Structure of the Hierarchical Earcons.

---

[*] In the electronic version of this paper, one can click on the nodes to hear the sounds.

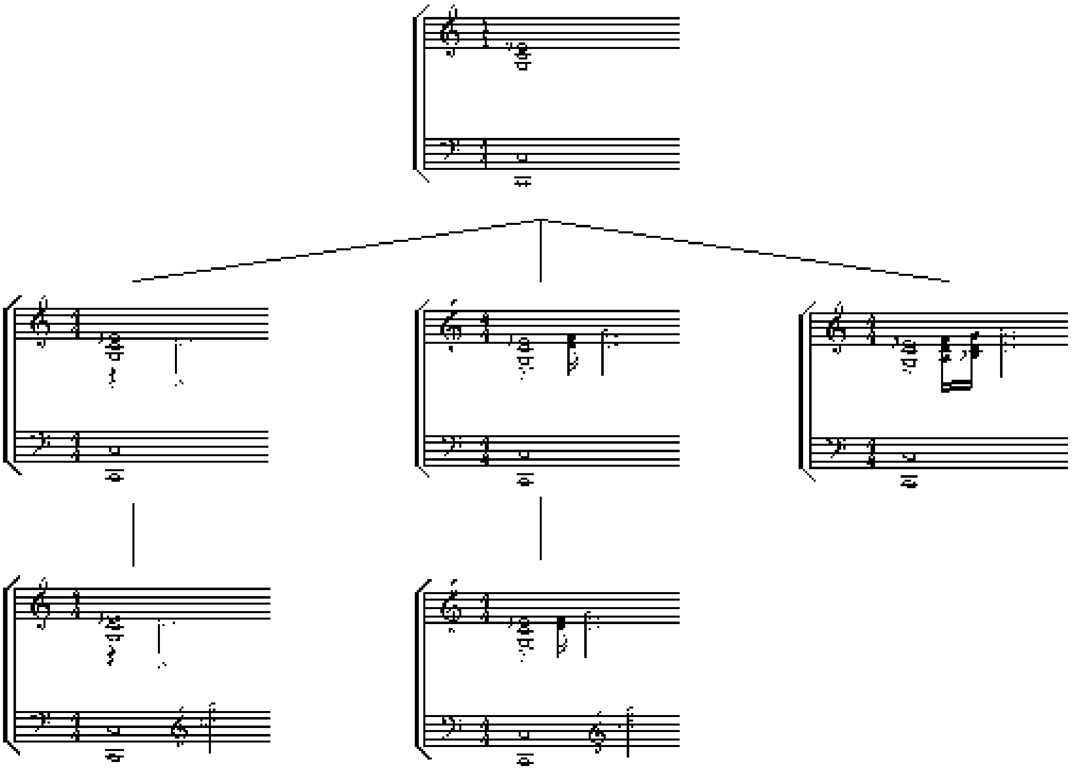

Figure 2: Family 1 sounds.

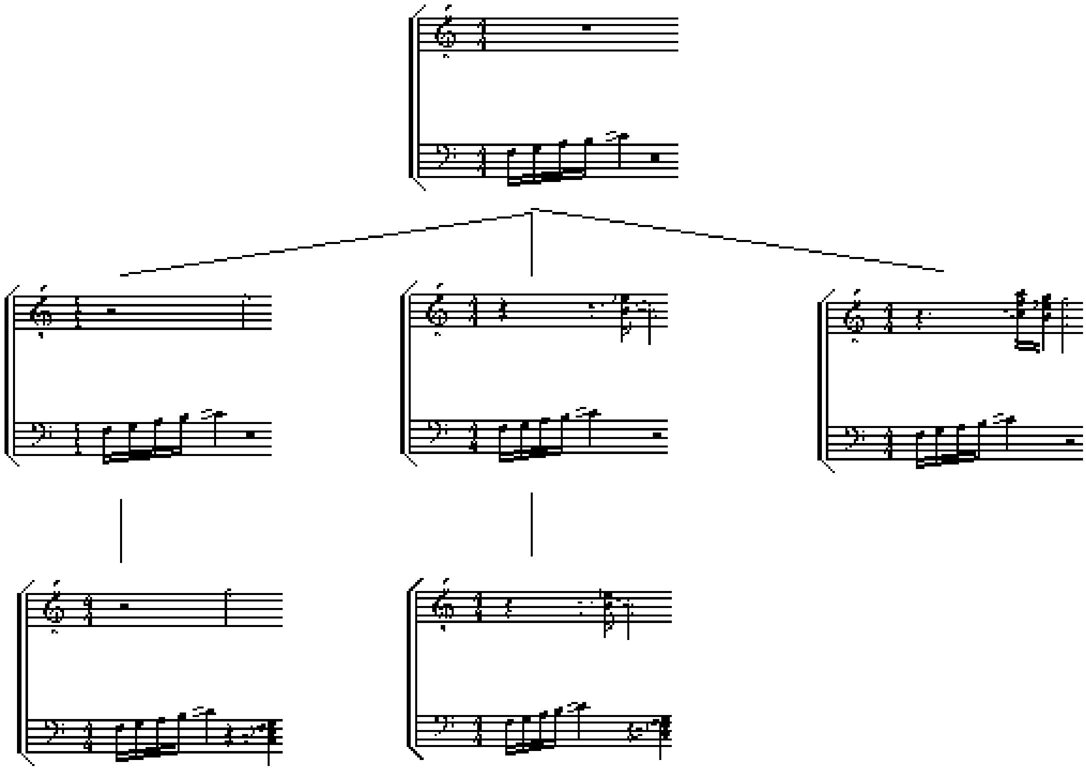

Figure 3: Family 2 Sounds.

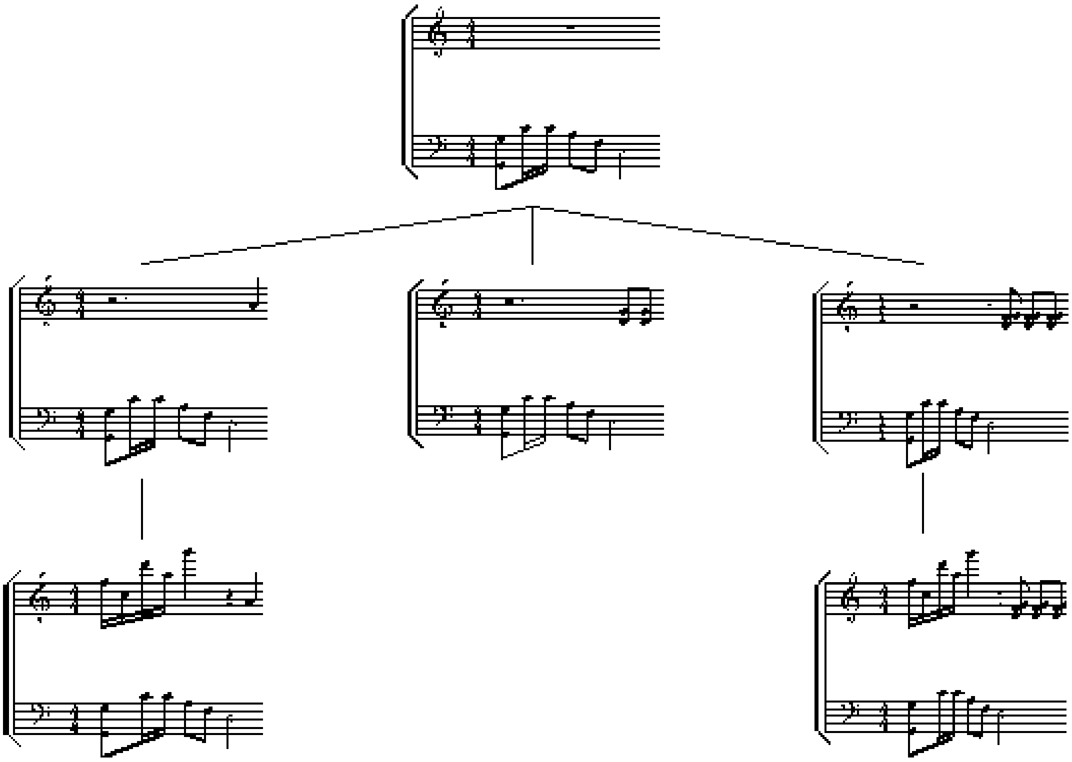

Figure 4: Family 3 Sounds.

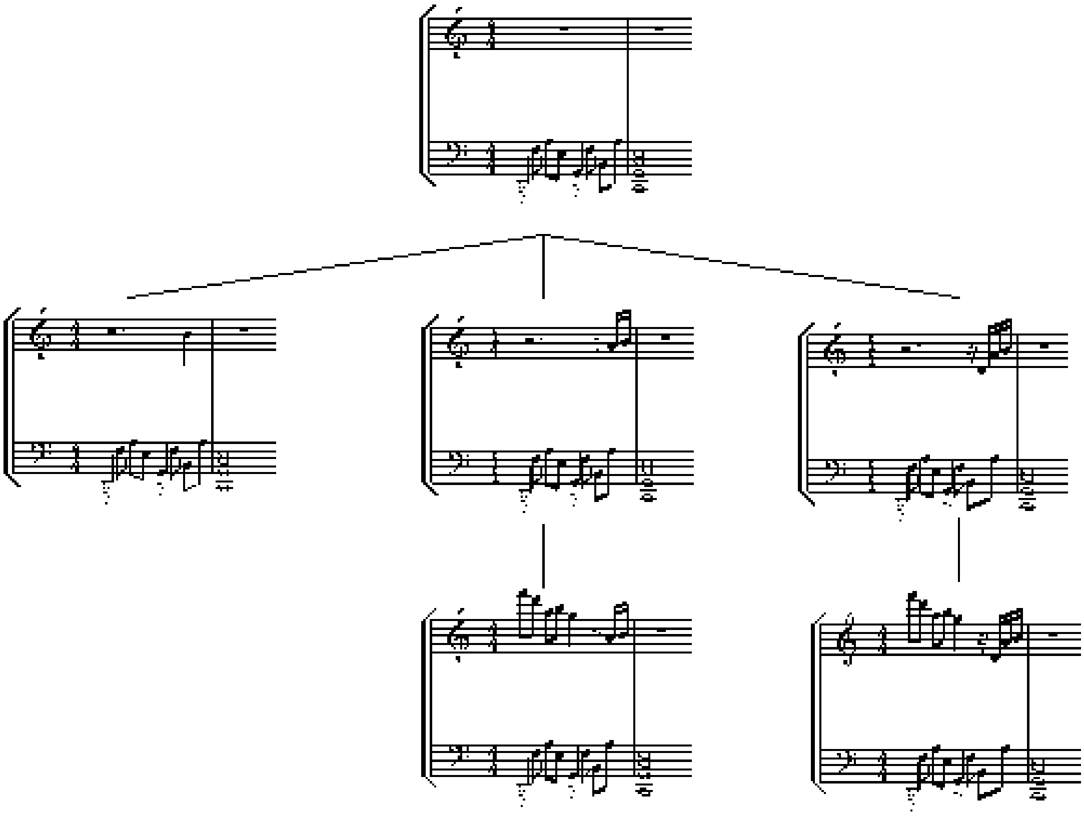

Figure 5: Family 4 Sounds.

3 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 [10] experiments for both musicians and non-musicians. This is why the same hierarchy was used. Besides, because we wanted to test other characteristics of the sounds, participants were asked to recall the 24 sounds of the hierarchy (all of them except the root) as opposed to 12 of the sounds in [7]. At the end of the experiment, participants were asked to fill a standard NASA TLX workload assessment form [19]. In this experiment, we also wanted to see how participants would manage simultaneous information within a non-trivial single instrument stream.

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 currently involved in musical activities.

3.1 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. Significantly better recall rates were expected for the earcons structured sequentially than for the parallel ones. No particular hypotheses were made about workload.

3.2 The Experimental Process

The design of the experiment was based on that given in [7] 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 [7]. 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 workload form on a graphical scale: Mental Demand, Physical Demand, Time Pressure, Effort Expended, Performance Level Achieved, Frustration Experienced, and Annoyance Experienced.

4 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 Student t-test showed that there was a significant difference between the results of the two groups (T_{10}=2.54, p=0.020).

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). There was no significant difference with the results of the first experiment described in [9] 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 [7] (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 3 (90.15%) then family two (84.1%) and lastly family 4 (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, 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 6 shows the recall rate of each node of the hierarchy

The workload tests performed on participants revealed that the sounds were not found annoying. We did not have any expectation concerning the annoyance or pleasantness of the sounds since they have been designed for an experimental purpose and are not meant to be used in an actual system. The mental demand and effort expanded were rated over 50% for the overall experiment, for both groups. Interestingly, all the items of the workload form (except "performance achieved") were rated quite similarly by all the participants, regardless their performance. "Performance achieved" matched the actual performance participants achieved, showing that they were aware of their understanding of the hierarchy. Some of them mentioned that the parallel earcons were demanding more effort to be recalled, but as we noticed above, the recall rates of sequential earcons were not significantly better than the recall rates of parallel earcons.

The recall rates for the third family (parallel earcons) were better than the recall rates for the second family (sequential earcons). Even if this difference is not significant, it is mostly unexpected since we Hypothesised a significantly higher recall rate for family 1 and 2 than for family 3 and 4. It is especially noticeable that the presupposed complex 1.3.3.1 sound was perfectly recalled (100% recall rate) by all the participants. Besides, it is hard to explain why the recall rate for the very similar 1.3.1.1 sound hardly achieves 70%. More generally, the irregularity of the results displayed in Figure 7 did not confirm our hypothesis concerning sequential versus parallel earcons.

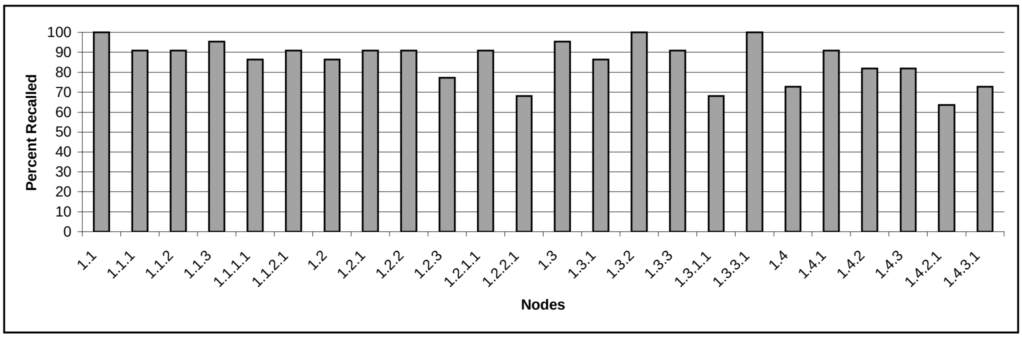

Figure 6: Overall Recall Rates of the 24 Nodes.

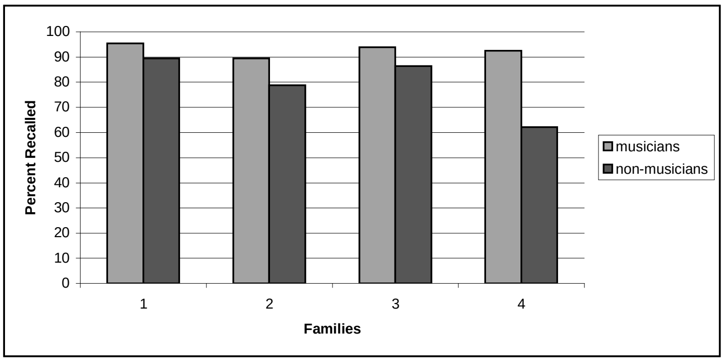

Figure 7: Recall Rates of the 4 Families for Musicians and Non-Musicians.

5 Discussion

We can advance certain arguments 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. 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 particularity of the third family was its fast tempo. Before the experiment, we were concerned about the fact that it would make the sounds harder to understand and recall, as the information they contain has to be processed more quickly. On the contrary, 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 let 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 find out the arguments that explain their performance. The actual task participants had to perform can roughly decomposed as follows:

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

These two tasks involve two different kind of mental activities. Remembering each of the families root sounds requires categorising correctly 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 4 root sounds, confusing the different families. We can explain this phenomenon 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 [10], 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 quite clear for level 4 sounds of the fourth family, the most complex of all. Again, we can presume that using different instruments will facilitate the separation of the relevant streams within each sound.

6 Conclusion

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. Therefore, the combination of this method with Brewster *et al.*'s design principles [10] should provide us with a more powerful auditory language of representation.

In the future, this language could be improved again by using the wide possibilities of sound synthesis to enhance the possibilities offered by musical semantics. The possibilities for conveying information without requiring users to have an analytical listening of the audio cues, e.g. by using the musical meanings of audio materials as much as possible, will be investigated further.

Since the objective of this research is to improve interaction in telephone-based interfaces, the analysis of actual navigational task problems in real-world system will determine the next step in the improvement of this auditory system of representation. Moreover, it will help us state about the ways these audio cues should fit within the design of the interface. According to the results of this task analysis, and with the design principles now available, we hope to build an effective hierarchy of sounds that would fit in a real-world system.

7 Acknowledgements

This research was founded by EPSRC; grant number GR/L66373.

8 Reference List

1. Alty, J. L. Can we use music in human-computer interaction? In Proceedings of HCI'95, Cambridge University Press, 1995, pages 409-423.
2. Alty, J. L., and Vickers, P. The caitlin auralization system: Hierarchical leitmotif design as a clue to program comprehension. In Proceedings of ICAD'97, Palo Alto, CA, USA, 1997, pages 89-95.
3. Alty, J. L., Vickers, P., and Rigas, D. Using music as a communication medium. In Extended Abstracts Proceedings of CHI'97 conference on human factors in computing systems, Atlanta, GA, ACM Press, 1997, pages 30-31.
4. Blattner, M. M., and Greenberg, R. M. Communicating and learning through non-speech audio. In Edwards, A. and Holland, S., (Eds.), Multimedia Interface Design in Education, Springer-Verlag, 1992, pages 133-143.
5. Blattner, M. M., Sumikawa, D. A., and Greenberg, R. M. Earcons and icons: Their common design principles. Human-Computer Interaction, 1989, 4(1): 11-44.
6. Blattner, M. M., Papp, A. L. III, and Glinert, E. P. Sonic enhancement of two-dimensional graphics displays, In Kramer, G. (Ed.), Auditory Display. Addison-Wesley, 1994, pages 447-470.
7. Brewster, S.A. Navigating telephone-based interfaces with earcons. In Proceedings of BCS HCI'97 (Bristol, UK), Springer Verlag, 1997, pp 39-56.
8. Brewster, S. A., Raty, V.-P., 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, pages 222-227.
9. Brewster, S. A., Leplâtre, G., Crease, M. G. Using Non-Speech Sounds in Mobile Computing Devices.

In Johnson C. (Ed.) Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices. Department of Computing Science, University of Glasgow, 1998, pages 26-29.

10. Brewster, S. A., Raty, V.-P., and Kotekangas, A. Earcons as a method of providing navigational cues in a menu hierarchy. In Proceedings of HCI'96 (Imperial College, London, UK), Springer-Verlag, 1996, pages169-183.
11. Brewster, S. A., Wright, P. C., and Edwards, A. D. N. Experimentally derived guidelines for the creation of earcons. In Adjunct Proceedings of BCS HCI'95, Huddersfield, UK, 1995, pages 155-159.
12. Cook, N. Perception: a perspective from music theory. In Rita Aiello with John Sloboda (Eds.), Musical Perception. Oxford University Press, 1994.
13. Gaver, W. W. Auditory interfaces. In M. G. Helander, T. K. L. Landauer and P. Prabhu (Eds.) Handbook of Human-Computer Interaction. Amsterdam: Elsevier Scienc . Second Edition, 1997.
14. Helmreich, R. Ergonomic user interfaces on telephones – are they worth it? In Proceedings of the 16th International Symposium on Human Factors in Telecommunications. Oslo, Norway, 1997, pages 15-20.
15. Leplâtre, G, and Brewster, S. A. Perspectives for the design of musical auditory interfaces. In Daniel Dubois (Ed.), Proceedings of CASYS'98 (Liège, Belgium), CHAOS asbl, 1998.
16. Lodha, S. K., Beahan, J., Heppe, T., Joseph, A., and Zane-Ulman, B. Muse: A musical data sonification toolkit. In Proceedings of ICAD'97, Palo Alto, CA, USA, 1997, pages 61-64.
17. Maguire, M. A human-factors study of telephone developments and convergence. Contemporary Ergonomics, 1996, pages 446-451.
18. Marx, M., and Schmandt, C. Mailcall: Message presentation and navigation in a nonvisual environment. In Proceedings of CHI'96 conference on human factors in computing systems, ACM Press, 1996, pages 165-172.
19. NASA Human Performance Research Group. Task Load Index (NASA-TLX) v1.0 computerised version. NASA Ames Research Centre, 1987
20. Roberts, T. L., and Engelbeck, G. The effects of device technology on the usability of advanced telephone functions. In Proceedings of CHI'89 conference on human factors in computing systems, ACM Press, 1989, pages 331-337.
21. Rosson, M. B. Using synthetic speech for remote access to information. Behaviour Research Methods, Instruments and Computers. 1985, 17(2):250-252.
22. Schmandt, C. Phoneshell: the Telephone as Computer Terminal. In Proceedings of the conference on Mulimedia'93. 1993, pages 373-382.
23. Smoliar, W. The role of music in multimedia. IEEE Multimedia, 1994, pages 9-11.
24. Schumacher, R. M., Hardzinski, M. L., and Schwartz, A. L. Increasing the usability of interactive voice response systems. Human Factors, 1995, 57(2):251-254.
25. Williams, D., Cheepen, C., and Gilbert, N. Experiments in how automated systems should talk to users. In Proceedings of HCI'98 Short Papers, 1998, pages 44-45.
26. Wolf, C., Koved, L., and Kunzinger, E Ubiquitous mail: speech and graphical interfaces to an integrated voice/email mailbox. In Proccedings of Interact'95 (Lillehammer, Norway), Chapman & Hall, 1995, pages 247-252.
27. Yankelovich, N., Levow, G. and Marx, M. Designing SpeechActs: Issue in speech user interfaces. In Proceedings of CHI'95 conference on human factors in computing systems. ACM Press, Addison-Wesley, 1995, pages 369-376.
28. Zajicek, M., and Brownsey, K. Methods for traversing a pre-recorded speech message network to optimise dialogue in telephone answering systems. In Proceedings of Eurospeech'93, Berlin, 1993.
29. Zajicek, M., and Powell, C. Accessing the world wide web by telephone. In Proceedings of the 16th International Symposium on Human Factors in Telecommunications, Oslo, Norway, 1997, pages 177-184.