Using Non-Speech Sounds in Mobile Computing Devices

Stephen Brewster, Grégory Leplâtre and Murray Crease

Glasgow Interactive Systems Group
Department of Computing Science, University of Glasgow, Glasgow, G12 8QQ, UK.
Tel: +44 (0)141 330 4966
Fax: +44 (0)141 330 4913
Email: stephen/gregory/murray@dcs.gla.ac.uk
Web: www.dcs.gla.ac.uk/~stephen/

INTRODUCTION

One of the main problems with output from small, hand-held mobile computing devices is the lack of screen space. As the device must be small to fit into the user's hand or pocket there is no space for a large screen. Much of the work on presentation in standard, desktop interfaces relies on a large, high-resolution screen. The whole desktop metaphor, in fact, relies on such a method of presentation. This means that much of the research on effective screen design and information output cannot be generalised to mobile devices. This has resulted in devices that are hard to use, with small text that is hard to read, cramped graphics and little contextual information.

Lack of screen space is not a problem that can easily be improved with technological advances; the screen must fit on the device and the device must be small; screen space will always be in short supply. Another problem is that whatever screen there is will be unusable in a mobile telephone once the device is put to the user's ear to make or receive a call.

There is one output channel that has, as yet, been little used to improve interaction in mobile devices (in fact very few systems of any type have used it effectively): Sound. Speech sounds are of course used in mobile phones when calls are being made but are not used by the telephone to aid the interaction with the device. Non-speech sounds are used for ringing tones or alarms (often in a quite sophisticated way) but again do not help the user interact with the system. There is now considerable evidence to suggest that sound can improve interaction [2, 4, 13] and may be very powerful in limited display devices.

We suggest that sound, particularly non-speech sound can be used to overcome some of the limitations due to the lack of screen space. Non-speech sounds have advantages over speech in that they are faster and language independent. Research we are undertaking has shown that using non-speech sound can significantly increase usability without the need for more screen space.

The rest of this position paper will outline some of the work we are doing with non-speech sounds to improve the usability of human-computer interfaces. More details can be found on the web site above.

Non-speech sounds

The sounds used in our research are all based around structured audio messages called *Earcons* [1, 5]. Earcons are abstract, musical sounds that can be used in structured combinations to create sound messages to represent parts of a human-computer interface. Detailed investigations of earcons by Brewster and colleagues [3, 9, 13] showed that they are an effective means of communicating information in sound. Earcons are also the most well-tested form of non-speech sounds.

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 (timbre, register, intensity, pitch and rhythm) *hierarchical* earcons can be created [1]. Using these techniques structured combinations of sounds can be created and varied in consistent ways. The earcons we have used are simple and within the range of sounds playable on a hand-held device. A set of guidelines to help interface designers use earcons are available [10].

Navigation in non-visual 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 voicemail) or interfaces to mobile telephones where the screen is too small to present the hierarchy of options available. In a telephone-based interface a user might call their 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.* [8] 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.

To solve this problem we have used non-speech sounds to provide navigation information [11, 14]. By manipulating the parameters of sound (timbre, pitch, rhythm, etc.) a hierarchy can be constructed. We constructed a hierarchy of four levels and 25 nodes with a sound for each node. The sounds we constructed based around rules – this meant that the users did not have to learn 25 different sounds but five simple rules. The idea then was that users would be able to hear a sound and from it work out their location in the hierarchy of options.

We have studied several different types of sounds and many different types of training and found very good recall rates. Users have been able to recall from 80% up to 97% of the sounds with minimal training and they could remember them over long periods of time. We believe that using sounds in this way can overcome many of the problems of presenting menus in small, hand-held devices. Users would also be able to move through the menus in a voicemail or telephone banking system, or the menu options within the telephone itself without becoming lost.

Sonically-enhanced widgets

Another important use of non-speech sounds has been to overcome problems of overloaded screens on desktop human-computer interfaces [13]. Graphical displays can become cluttered and hard to use when too much information is put onto them. This is especially true in mobile devices where screens are small. Using *sonically-enhanced* widgets (such as buttons, scrollbars, etc.) means that information can be moved off the graphical display and put into sound, thus freeing the display for the most important information.

Some desktop interfaces allow the use of sounds (e.g. Windows95) but these are often gimmicks rather than additions to improve usability. This means that they are often perceived as annoying. Our sounds were added to interface components to address specific usability problems. The sounds helped our users work more effectively and therefore they did not find them annoying (as demonstrated by our experimental results).

The results of our experiments on buttons, scrollbars, menus, drag and drop, and tool palettes have shown a significant improvement in usability: time taken to complete asks and time to recover from errors were reduced, subjective workload was reduced and user preference increased [6, 12, 15, 16]. In each case we also checked to see if annoyance had been increased by the addition of sound. In none of the widgets did users rate the sonically-enhanced interface more annoying than the standard, graphical one.

As a simple example, in one of our experiments we removed the visual highlight from graphical buttons and replaced it with sound. The sounds told users when they were on the button and when they had pressed it correctly (which can be hard to see and causes users to ‘slip-off’ a button and think it has been pressed when it has not [6]). The sounds overcame these basic interaction problems with buttons and increased usability. Users had no problem with the lack of graphical feedback (in fact they performed much better with the sonically-enhanced buttons).

We suggest that this could also allow widgets to be reduced in size, thus saving space on a limited graphical display but without compromising usability (if the widgets were simply reduced in size then they would become harder to use and more error-prone). In the case of the buttons described above, the sounds could tell users when they were on the button even if it was small and the tip of a pointing device was obscuring it. The sounds would confirm if the button had been pressed correctly, whereas the limited amount of visual feedback from the small graphical button would be easily missed.

Presenting dynamic information on hand-held devices

One potentially large group of mobile device users is people who need to access dynamic information on the move. For example, stock market traders who want to monitor share prices or systems administrators who want to monitor network performance when offsite. Mobile devices do not currently support such activities very well.

Presenting dynamic information on devices with small displays is difficult. Techniques for large displays again do not generalise well to small ones. For example, on a large graphical display a strip-chart or line graph might be used so that users can see a change in one or more variables over time. This is very effective as the user can see general trends, compare one graph against another and also find current values. Utilising this type of presentation on a small display device is impossible as the screen is too small to provide any history so that trends are difficult to follow. If the user has the device to his/her ear then no visual display is possible.

One way to present this type of information is using sound. Research has shown that presenting graphs in sound for blind people allows them to extract much of the information needed from the graph [7]. We are working on solutions based on non-speech sounds to present the trend information with speech to present current values when required.

We are also investigating more complex graphical visualisation techniques for example presenting complex hierarchical information (such as file systems) based on the navigation work described above. All of the research in this area is aimed at making mobile devices more useful to users by allowing them access to more complex information than they can currently get at.

CONCLUSIONS

A solution is needed to the problem of lack of screen space on mobile devices. This is unlikely to come from larger screens because the devices are physically limited in size to fit into the user's hand. We suggest that non-speech sounds can solve many of these problems without making the devices larger.

Structured non-speech sounds (in combination with speech) can provide information to help users navigate through menus of options more easily. Sounds can improve the usability of on-screen widgets so that screen clutter can be avoided. They can also provide access to more complex, dynamic information that is currently very difficult to use on hand-held devices.

REFERENCES

1. Blattner, M., Sumikawa, D. and Greenberg, R. Earcons and icons: Their structure and common design principles. *Human Computer Interaction* 4, 1 (1989), 11-44.
2. Gaver, W., Smith, R. and O'Shea, T. Effective sounds in complex systems: The ARKola simulation. In *Proceedings of ACM CHI'91* (New Orleans) ACM Press, Addison-Wesley, 1991, pp. 85-90.
3. 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.
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. 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 ACM/IFIP INTERCHI'93* (Amsterdam, Holland) ACM Press, Addison-Wesley, 1993, pp. 222-227.
6. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. The design and evaluation of an auditory-enhanced scrollbar. In *Proceedings of ACM CHI'94* (Boston, MA) ACM Press, Addison-Wesley, 1994, pp. 173-179.
7. Edwards, A.D.N., Pitt, I.J., Brewster, S.A. and Stevens, R.D. Multiple modalities in adapted interfaces. In *Extra-Ordinary Human-Computer Interaction*, Edwards, A.D.N. (Ed.), Cambridge University Press, Cambridge, UK, 1995, 221-243.
8. Yankelovich, N., Levow, G. and Marx, M. Designing SpeechActs: Issues in speech user interfaces. In *Proceedings of ACM CHI'95* (Denver, Colorado) ACM Press, Addison-Wesley, 1995, pp. 369-376.
9. 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.
10. 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, pp. 155-159.
11. Brewster, S.A., Raty, V.-P. and Kortekangas, A. Earcons as a method of providing navigational cues in a menu hierarchy. In *Proceedings of BCS HCI'96* (London, UK) Springer, 1996, pp. 169-183.
12. Brewster, S.A. and Crease, M.G. Making Menus Musical. In *Proceedings of IFIP Interact'97* (Sydney, Australia) Chapman & Hall, 1997, pp. 389-396.
13. Brewster, S.A. Using Non-Speech Sound to Overcome Information Overload. *Displays* 17 (1997), 179-189.
14. Brewster, S.A. Navigating telephone-based interfaces with earcons. In *Proceedings of BCS HCI'97* (Bristol, UK) Springer, 1997, pp. 39-56.
15. Brewster, S.A. The design of sonically-enhanced widgets. *Accepted for publication in Interacting with Computers* (1998).
16. Brewster, S.A. Using earcons to improve the usability of tool palettes. In *Summary proceedings of ACM CHI'98* (Los Angeles, Ca) ACM Press, Addison-Wesley, 1998, pp. 297-298.