



KUNGL  
TEKNISKA  
HÖGSKOLAN



CID-242 • ISSN 1403-0721 • Department of Numerical Analysis and Computer Science • KTH

## **Designing Accessible Auditory Drag and Drop**

**Winberg, F. & Hellström, S. O.**

**Proceedings of the 2003 Conference on Universal Usability, CUU2003.**

**Vancouver, Kanada**



**CID, CENTRE FOR USER ORIENTED IT DESIGN**

## **Winberg, F. & Hellström, S. O.**

Designing Accessible Auditory Drag and Drop

Proceedings of the 2003 Conference on Universal Usability, CUU2003. Vancouver, Kanada,

**Report number:** CID-242

**ISSN number:** ISSN 1403 - 0721 (print) 1403 - 073 X (Web/PDF)

**Publication date:** November 2003

**E-mail of author:** fredrikw@nada.kth.se

### **Reports can be ordered from:**

CID, Centre for User Oriented IT Design

NADA, Department of Numerical Analysis and Computer Science

KTH (Royal Institute of Technology)

SE- 100 44 Stockholm, Sweden

Telephone: + 46 (0)8 790 91 00

Fax: + 46 (0)8 790 90 99

E-mail: cid@nada.kth.se

URL: <http://cid.nada.kth.se>

# Designing Accessible Auditory Drag and Drop

Fredrik Winberg

Centre for User Oriented IT-Design  
Royal Institute of Technology  
SE-100 44 Stockholm, Sweden  
+46-8-7909272

fredrikw@nada.kth.se

Sten Olof Hellström

Centre for User Oriented IT-Design  
Royal Institute of Technology  
SE-100 44 Stockholm, Sweden  
+46-8-7906698

soh@nada.kth.se

## ABSTRACT

This paper presents an audio-only version of drag and drop. By continuously presenting the information, using auditory zooming at two different levels and absolute positioning of the cursor, a blind user is able to get an overview, locate and interact with a specific object. Two user studies on two different versions have been made in order to get input to the design process and to evaluate the ideas. The results points at the importance of being able to customize the interface and to provide an overview of all interface objects.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces – auditory (non-speech) feedback, interaction styles; K.4.2 [Computers and Society]: Social Issues – assistive technologies for persons with disabilities;

## General Terms

Design, Human Factors

## Keywords

Auditory interface, accessibility, blind users, drag and drop.

## 1. INTRODUCTION

This work is about investigating new techniques for giving blind users better access to graphical user interfaces using sound. The major difference between screen reading software for blind computer users and ordinary graphical user interfaces is the difference in presentation of the information (cf. [2]). The screen reader presents the contents of the screen in a line-by-line fashion, using speech synthesis or Braille. This linear presentation does not allow for presentation of concurrent and spatial information in the same way as a graphical user interface does.

### 1.1 Previous research

Ways of presenting graphical user interfaces for blind computer users has been explored a number of times before. In the Mercator project [3], the graphical user interface was presented using a hierarchical model of the interface objects

where the logical relationship is represented using everyday sounds. In the GUIB project [4], the spatial relations of the objects are presented using a novel tactile device.

Building on the experience from earlier studies on auditory direct manipulation [6], a sonification model has been designed that implements drag and drop, and that has the same resolution as a computer screen.

## 1. AUDITORY INTERFACE

The two versions of the auditory interface described here was designed to support drag and drop, which involves movement of objects by positioning a pointer on the object to be moved, picking it up, dragging it to the desired location and dropping it there. In order to do this, the interface must support getting an overview of all objects, locating a specific object, and interaction with the object.

This is accomplished by using auditory zooming, a technique where the granularity of the auditory display increases when the user “moves closer” to the information, mapping many objects to one sound when far away, and one object to one sound when close [1]. In this implementation, the presentation is divided into two levels, the *overview* and the *zoomed view*.

### 1.1 Overview

The overview gives the user a general notion of where there are objects, and approximately how many. This is accomplished by dividing the screen into four quadrants. The location of the objects is presented using four different tones. The tones are separated using pitch (high or low representing top and bottom) and stereo panning (left or right representing left and right). For each object added to a quadrant, overtones are added to the tone, representing an approximation of the number of objects in that quadrant. In the first version the tones are repeated in parallel continuously. In the second version the high and low tones altered between left and right separately, creating two parallel sequences.

### 1.1 Zoomed view

The zoomed view gives the user detailed information about a subset of the display, the quadrant in which the pointer is located. All objects placed in the same quadrant as the pointer are audible. The volume of each object depends on the distance to the pointer, the closer an object is, the higher the volume. The objects are presented one by one.

### 1.1 Objects

In the first version, all objects were based on the same sound. In the second, the objects had separate sounds. The sound changed depending on where the object was located with respect to the pointer. In the first version, the vertical location was represented by the length of the sound (long, short or

Copyright is held by the author/owner(s).

CUU'03, November 10-11, 2003, Vancouver, British Columbia, Canada.

ACM 1-58113-701-X/03/0011.

middle representing whether the object is located below, above or at the same vertical level as the pointer). In the second version, an auditory cue was added to the sound (a high, low or middle pitched tone representing above, below or at the same vertical level). In both versions, the horizontal location was represented using stereo panning (left, right or middle representing left, right or the same horizontal level).

## 1.2 Manipulation

The user interacts with the objects using a pen stylus on a graphics tablet. This is used in order to have absolute positioning of the pointer, as opposed to the mouse whose relative positioning makes it harder to use sound as the only output device when the complexity of the display is large (cf. [5]). Additionally, using the mouse requires sonification of the position of the cursor, which limits the auditory bandwidth left for sonifying other components (cf. [6]).

There are also event driven sounds that gives the user feedback on specific actions, in order to emphasize the directness and physical nature of the interaction. These includes when hitting, picking up, dragging, and dropping an object.

## 1.3 Manipulating auditory parameters

In order to give the user more control over the auditory environment, the second version provided the user with a MIDI controller device with a number of sliders. The first slider was a cross-fade between the overview presentation and the objects. The second controlled the volume of the added tones that represented vertical location and distance. The third controlled the overall volume. By providing this kind of device, the user would be able to adapt the display interactively and continuously while exploring and interacting.

## 2. USER STUDIES

The main object of the two user studies was to get input from the user group in an early stage of the design process. This means that rather than taking quantitative measurements of for example task completion time, concrete design ideas and suggestions for alterations was the main focus. The second version is not to be seen as an improved or better version, but rather, but rather a tool to try out a number of design ideas that came up during the first user study.

Each subject was introduced to the auditory interface and was encouraged to just play around with the interface. They were asked to perform a few tasks, such as counting the number of objects on the screen or moving all objects to a specified part of the screen.

Four blind subjects participated in the first study, and three of these also participated in the second study. This gave the subjects a possibility to see how their design ideas from the first study were being carried out and how they worked.

## 3. RESULTS

Even though the subjects were initially overwhelmed by the complexity of the sound environment initially, it did not take long for each of them to learn how to locate and move the objects around.

When having a complex auditory environment it is important to provide the users with means of individualizing the display. This is important not only in terms of supporting different user needs and expectations, but also to support different tasks. Even when having very few subjects as in the studies presented here, we observed different ways of interacting with the auditory interface, especially the use of the MIDI controller and how the exploration with the pointer was performed. In order to do this, it is important to provide the user with different ways of obtaining the same information. In the auditory interface described in this paper, there were for example two different ways of finding out where the objects were located on the screen, either using the overview presentation or positioning the pointer in one quadrant after another.

Despite the fact that the subjects hardly did use overview presentation except when asked for, we do still believe that this is an important part of an auditory interface. A long term effect of the overview presentation would preferably be that the user would not have to pay any special attention to it, but rather be given a general sense, a gestalt, of the distribution of objects on the screen. Additionally, providing a continuous presentation of all objects does not require the user to leave the current position and perhaps loose the focus of the object being interacted with. Even if the overview presentation was not used in an apparent way in these studies, the subjects did agree that if given time the overview presentation could have the effect that we did intend. However, when asked, the subjects did succeed in interpreting this presentation.

## 4. REFERENCES

- [1] Axen, U. & Choi, I. Investigating geometric data with sound. in Proceedings of the Third International Conference on Auditory Display ICAD'96.
- [2] Boyd, L.H., Boyd, W.L., and Vanderheiden, G.C. The graphical user interface: Crisis, danger and opportunity. Journal of Visual Impairment and Blindness, (December 1990), 496-502.
- [3] Edwards, W.K., Mynatt, E.D., and Stockton, K. Providing Access to Graphical User Interfaces — Not Graphical Screens. in Proceedings of the first annual ACM conference on Assistive technologies, ASSETS'94, ACM Press, 47-54.
- [4] GUIB Consortium. Final Report of the GUIB Project: Textual and Graphical Interfaces for Blind People. Royal National Institute for the Blind, London, 1995.
- [5] Pitt, I.J., and Edwards, A.D.N. Pointing in an Auditory Interface for Blind Users. in IEEE International Conference on Systems, Man and Cybernetics 1995, 280-285.
- [6] Winberg, F., and Hellström, S.O. Qualitative aspects of auditory direct manipulation: a case study of the Towers of Hanoi. in Proceedings of the 7th International Conference on Auditory Display, ICAD 2001, 16-20).