Adapting Graphical User Interfaces for Use by Visually Handicapped Computer Users: Current Results and Continuing Research

Arthur I. Karshmer, Bill Ogden [‡]
Pres Brawner [†], *Karlis Kaugars & George Reiswig* [†]

Computer Science Department
Computing Research Laboratory [‡]
Psychology Department[†]
New Mexico State University
Las Cruces, NM 88003 USA
karshmer@cs.nmsu.edu

ABSTRACT. The use of modern computers and software by the visually handicapped has become more difficult over the past few years. In earlier systems the user interface was a simple character based environment. In those systems, simple devices like screen readers, braille output and speech synthesizers were effective. Current systems now run Graphical User Interfaces (GUIs) which have rendered these simple aids almost useless. In no area has this problem become more important than in technologies for the handicapped. What has become enabling technology for the sighted has become disabling technology for the visually impaired. In the current work we discuss new and innovative approaches to permit non-sighted users to interface with GUIs, having the salutary effect of gaining needed access to the most modern computing equipment for a subset of our population that is otherwise excluded from such access.

Using our approach to integrating special interfaces into those already shipped by the manufacturer, the non-sighted user will no longer be isolated from the mainstream of the information world. The net effect of such an interface is to make computing and information resources available to the visually handicapped user at the cost as to the sighted user.

1 Introduction

The need to lend support to visually handicapped computer users has been acknowledged for some time. With the passage of laws such as the Americans with Disabilities Act and similar legislation in other countries, new attention has been focused on a serious societal problem. Delivering new and innovative technologies to visually handicapped students has become a high priority in our society. If attention is not paid to this problem, we are likely to create an underprivileged class in our information based world, not based on the normal delineations between rich and poor, but based on access to modern technologies.

A number of systems have been designed and developed over the years and a few have even made it to market. In somewhat typical projects [31, 30, 15, 16] the philosophical approach taken was to redesign the computing environment's hardware to suit the visually handicapped user. These systems involved the synthesis of technologies uniquely suited to the visually handicapped to provide a more *comfortable* interface without much concern for the bandwidth or the generality of the interface. As a result, such interfaces were suited only for use by the visually impaired. Recent thought has lead us to believe that this was a philosophically incorrect approach. Our current approach is to use a variety of highly organized sounds to enhance, but not alter, already existing visually based interfaces so they will provide a compatible and reasonably high bandwidth communication channel for sighted and non-sighted users on the same hardware/software platform. Thus, our new approach should lead to more economical products for the visually impaired, facilitate interactions between sighted and non-sighted users and enhance the delivery of informational material to the visually impaired.

The Graphical User Interface (GUI) has become the standard in virtually all computing environments available on a broad range of hardware/software platforms available in the domain we call "personal computers." The traditional character based interface has become the exception rather than the norm. From PCs to high-speed workstations to X Terminals, the "dumb terminal" interface is becoming history. While this trend enhances the interface for the sighted, it is yet another stumbling block for the visually impaired. The goal of our current research is to make the GUI available to the non-sighted user.

Recent developments in computer interfaces have made it possible to present ever increasing amounts of information to the user in a visually meaningful form. Window-based and graphical display systems make it possible to display in many windows, and in many forms, large amounts of information. Nearly all of this information is visually presented; only a small fraction is presented via auditory channels. With the rapid advances in, and availability of technology, sound enhanced user interfaces are becoming increasing possible and attractive. Research in many areas has produced results that can be directly applied to the problems of sound enhanced interface design.

Research into computer generated **virtual worlds** [5, 19] has produced the notion of an "audio window" in which a spatial sound presentation system positions sound sources associated with a three-dimensional audio space. "Filtears" are used to emphasize or mute particular audio windows based upon their current activity. Sound generation uses audio signal processing to exploit the psychoacoustic properties of human hearing. The generalization of display systems for three-dimensional acoustic information has been examined in [29], thus providing a reasonable expectation that non-individualized sound based interfaces can be constructed.

In the area of speaking interfaces, the XSpeak system [27] demonstrates that spoken commands can be used to aid in the navigation through a set of windows. Speech input proved to be neither faster nor slower than navigation with the mouse. Playback of digitally recorded speech for fixed text has long been possible; however, high quality text-to-speech synthesis for arbitrary text is also now possible [25].

Considerable research has been done to identify non-speech sounds for use in sound enhanced interfaces [11, 12, 7, 8]. In the Karshmer system, non-speech based sounds are being tested as tools to aid in the navigation of complex hierarchical structures: namely menus. The work by Edwards examined associating sounds with screen objects. Another proposal is to use a system of naturally occurring sounds or 'everyday sounds' to represent actions within the interface [9]. In another interesting work, a musical sound was used to represent the peaks and valleys in the infrared spectrum of a chemical compound [24]. It has also been demonstrated that a rich set of musical variables (rhythm, pitch, timbre, register, dynamics) can be used to establish families of the audio tones equivalent to *icons* [1] [1]. An excellent introduction to this area of research is given in [4].

While researchers have used sound to enhance learning and usability of computer interfaces, the choice of which sounds (pitches) to use has often been made without the benefit of solid empirical data. The frequent implication is that an individual who is not musically trained might not be able to make use of real musical information presented by the interface. There is however, a growing body of research which indicates that even untrained listeners can consistently identify sounds in a given tonal context. Studies have shown positive results in "well-fittedness," indicating the presence of perceivable qualities in sounds from an established context [17, 3].

Further, parallels have been shown between speech and sound cognition [6, 18], with results pointing to the possibility of developing a functional hierarchy of sounds to be attached to an interface/menuing system. The important finding here is that a listener's un-

[1] The term *earcon* has been suggested to be the auditory equivalent to the now commonplace icon [15].

derstanding of a particular aural event depends heavily on how that event relates to any prior events, and this in turn affects potential future events as well.

The implication for our work is that many of these theories have not yet been fully integrated, explored and tested empirically. We are developing hierarchical sets of context related sounds and attaching them to related features in an interface. By measuring learning and response times for non-sighted users, we expect to show that the appropriate use of such sounds can help visually impaired users learn a system faster and use it both faster and more efficiently than with more conventional systems. The major departure here is that we are embedding our enhancements into already existing visual interfaces (operating systems). The net result of the work will be to extend a common computing environment to a number of groups of users: each having different needs. The interface tools will present no obstacle to sighted users, and indeed may prove to be beneficial to them as well. Our approach will make the delivery of *special* services to *special* users transparent and therefore encourage major manufacturers to move in this necessary direction. In a world of different interfaces for different people, the result was predictable - if the special user group did not represent a significant number of potential sales, the probability of offering such products was dramatically reduced. Hence, special needs have often gone unmet by commercial developers.

In recent work [11, 12, 13] two novel approaches to presenting sound-based tools to visually handicapped computer users were described. One was based on the concept of navigating complex hierarchical menu structures using musical tones. This approach was based in part on work reported by Paap [22, 23]. The second involved the use of directional sound to aid the user in navigating through two- and three-dimensional spaces such as found in the typical GUI. The goal of this work is to demonstrate the feasibility of making the GUI usable by the visually handicapped user.

The current work describes a set of experiments that we have designed to test the ability of users to navigate the graphical *desktop* and hierarchical menu structures using sound. In the remainder of this work we describe our test environment and the methods we will use to gather and analyze the data we need to evaluate our work.

2 Problem Statement & Rationale

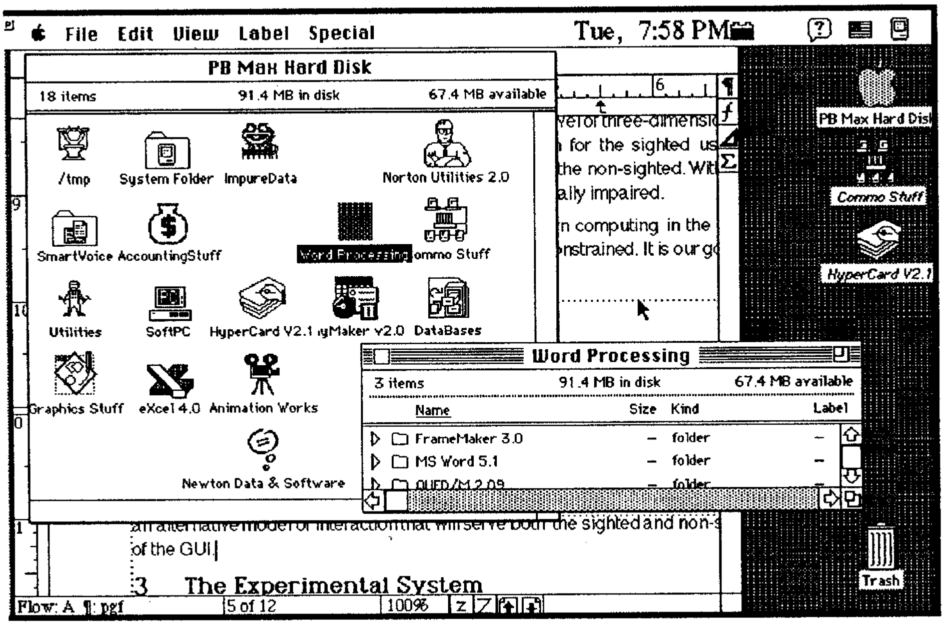
The first level interface encountered by the user of a typical GUI (for example the Macintosh) is the what is called the *desktop* (see Figure 1). This is a very common metaphor used in modern interfaces. The desktop is composed of several graphical objects, which can be freely moved on the screen, including:

- A Menu Bar
- Several systems icons on the menu bar (the Apple Menu, the Application Selector, Balloon Help Icon, etc.)
- Various open windows containing icons and/or lists.
- Various icons

While this representation seems quite natural to the sighted user, it presents the equivalent of a mine field to the non-sighted user. Objects on the screen can partially or completely occlude other objects, giving a level of three-dimensionality to the interface. While fully occluded objects are a problem for the sighted user, partially occluded objects further complicate the use of GUIs by the non-sighted. Without special help, such an interface is effectively useless to the visually impaired.

Given the prevalence of GUIs, the use of modern computing in the social environment by the visually handicapped is extremely constrained. It is our goal to design an alternative model of interaction that will serve both the sighted and non-sighted user of the GUI.

To this end we have designed a model of interaction based on the standard GUI which allows the visually handicapped user to navigate the desktop using sound-based information. The desktop is identical to the one used by sighted users, only the navigational cues



Word Processing

3 items 91.4 MB in disk 67.4 MB available

Name	Size	Kind	Label
FrameMaker 3.0	–	folder	–
MS Word 5.1	–	folder	–
QUED/M 2.09	–	folder	–

Figure 1 - A Typical GUI Desktop

are different. Navigating through the desktop is the prerequisite step to launching applications needed in the computational process. If the desktop is not accessible, then the underlying applications and other tools will not be accessible.

Generally, we have designed three basic tools to implement and test in order to achieve our goals. These will be described in greater detail in a later section of this work. Briefly stated, they are:

- Navigation of the desktop using a “sound map” and cursor warping[2]
- Navigation of the desktop using sonar vectoring of the cursor
- Sound-based menu navigation.

Given the similar nature of application interfaces to the desktop, our tools will be useful at the application level as well.

2.1 Navigation with a Sound Map

In this system, all items on the desktop are assigned some unique sound which can be selected by the user or by the system. If selected by the system, a combination of speech synthesis and sound generation is used to train the user. At any given instant, the user can request a sound based map of the desktop. As the appropriate sounds are generated by the system, the user on a click of the mouse can be instantly warped to the object just described. Once warped to the proper location, normal mouse click sequences will complete the user’s task. The soundmap approach should provide a facile interface with small amounts of training. The interesting questions concerning this interface relate to the use of appropriate families of sounds.

2.2 Navigation with Sonar Vectoring

Consider the desktop as a multi-dimensional collection of various objects, some in front of others. Further, consider the cursor as a sonar emitter and the user as the sonar receiver.

[2] Warping is the action of placing the screen cursor at a location selected by the system rather than directly by the user. While manufacturers such as Apple frown on this practice, it is critical in special purpose interfaces. Using warping in the special purpose interface has no negative impact on the normal interface.

At any instant the user can receive tonally based vectoring information to indicate the location of the cursor on the desktop and in relation to the various objects on the desktop. The user has the choice of continuous information or periodic reporting. The sonar "ping" returns a variety of information on each report:

- The general location of the cursor on the desktop. For example in which quadrant the cursor is located. Using the stereophonic output capabilities of modern systems, this two-dimensional information can be easily announced to the user.
- The object(s) under the cursor. If multiple objects, the report will indicate both the objects and their depth.
- The objects on or near the current path of the cursor if it is moving. By using a Doppler effect in the ping response, proximity information can be passed to the user.

Once the user identifies that (s)he is on the proper vector, a series of mouse clicks or simple key presses will warp the user to the desired object. Again, at this point standard mouse clicks will do the rest of the job.

2.3 Sound Based Menu Navigation

Given general desktop navigation, one key element needs to be developed to complete the set of tools for the visually impaired user: navigating and manipulating menus. Work in this area is currently underway in the research group. A model of multi-level menu systems has been implemented using HyperCard 2.1[3] which is currently in early beta testing. The model presents up to a six-level menu structure with up to six items on each level. For initial testing, we have restricted the structure to three levels.

Our basic premise is simply "**by use of appropriate tonal cues, computer users will be able to navigate through complex hierarchical menu structures which are organized in an identical manner to those used by sighted users**." When the final results are in, we expect to answer several practical and one fundamental question. Briefly stated the practical questions are:

- Can menu structures be navigated using verbal cues?
- Can menu structures be navigated using tonal cues?
- What is the difference between the ability of users to navigate such structures using the different clues?
- Do truly non-sighted users have a better ability to navigate these structures than blindfolded sighted users?
- What does all of this imply about how sighted vs. non-sighted people represent and use navigational information?

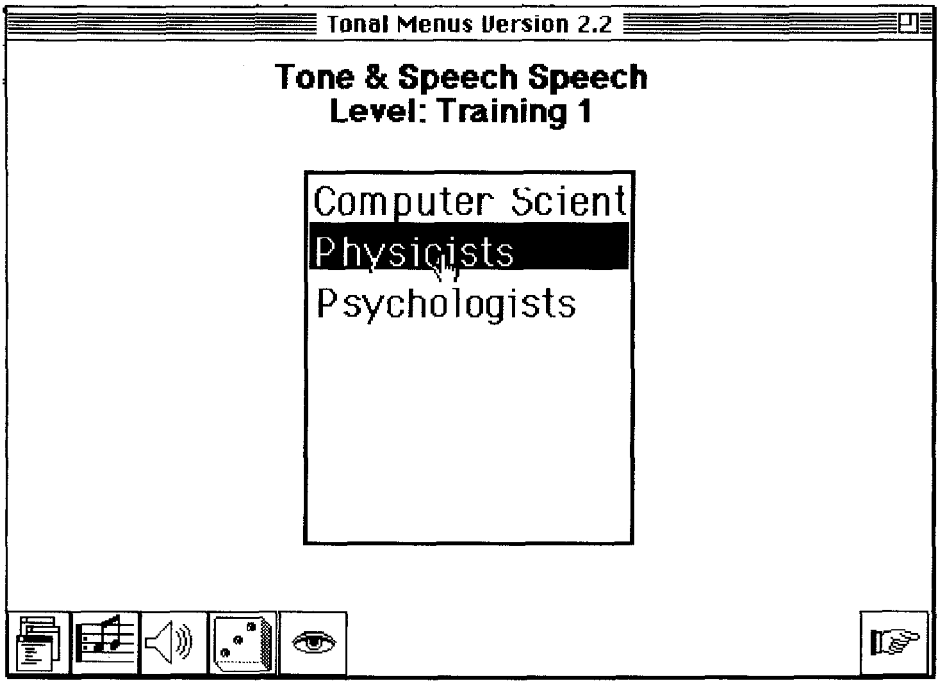
To test this hypothesis we have developed a test environment that allows us to collect data on a large number of subjects. The subjects fall into two major groups: blindfolded, but sighted, students here at the university and truly non-sighted students at a nearby school for the visually handicapped. Both groups are presented with the same interface.

The interface was developed using the HyperTalk language and runs on any member of the Macintosh family of computers. Snapshots of two levels of the experimental interface is shown in the following figures.

While the menu's structure appears on the screen, the test subject will not be able to see it as (s)he is either blindfolded or truly visually impaired. The control buttons in the lower left corner are used by the experimenter to configure the system. These buttons are disabled once the test session begins.

The subjects are automatically positioned at the top entry (home position) of the menu structure at each level. When given a path to follow they move the mouse up or down

[3] HyperCard, and its programming language HyperTalk, is an extremely powerful rapid prototyping tool distributed by Apple Computer Corp. Building interface tools with HyperCard allows us to design and test interface concepts in less than 1/10th of the time needed using other programming tools.

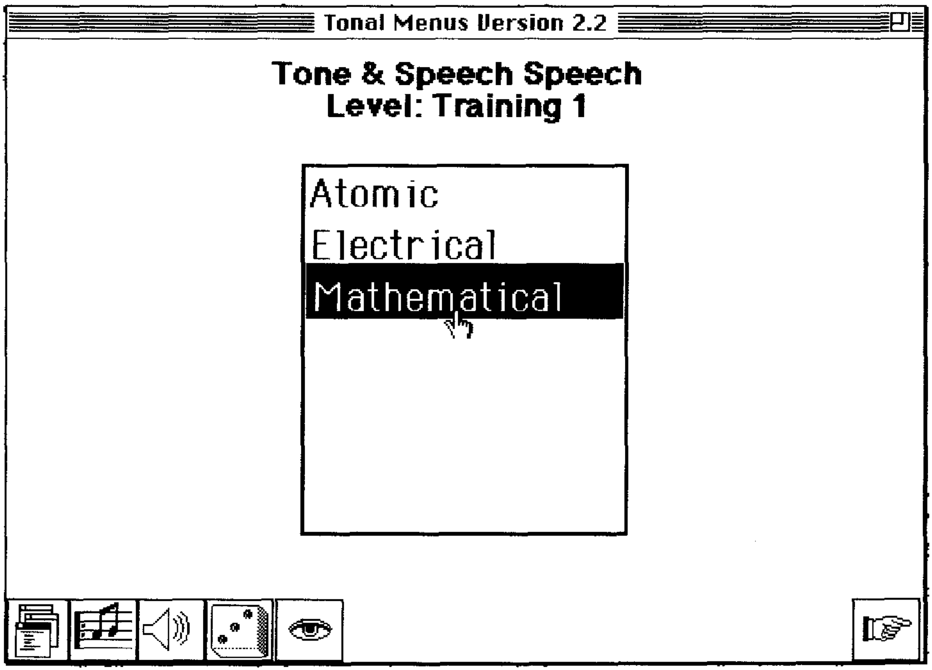


Tonal Menus Version 2.2

Tone & Speech Speech
Level: Training 1

Computer Scient
Physicists
Psychologists

Figure 2 - The First Level of the Test Menu



Tonal Menus Version 2.2

Tone & Speech Speech
Level: Training 1

Atomic
Electrical
Mathematical

Figure 3 - The Second Level of the Test Menu

receiving auditory cues (either tonal, voice, or both). At any point in the search the user can move down a level by a single click on the mouse. Further, the test subject can request information on the current location of the cursor by clicking with the shift-key held down. Once the test path is given, data are recorded in the areas of total mouse clicks to goal, time to goal and number of help requests made by the subject. In our initial tests, paths are restricted to a width of three and a depth of three.

Two methods have been implemented to help the user keep the cursor within the bounds of the menu. First, if the user strays out of bounds, a special auditory signal informs the subject where (s)he is located in relation to the menu. Second, we have the ability to constrain the movement of the cursor in either the horizontal or vertical direction. Through the use of specially developed HyperTalk XCMDs and XFCNs[4], we are able to ***warp*** and restrict cursor motion.

The tones generated in each category, each level and each item at each level have been designed to pass on the maximum information concerning the user's location in the hierarchy. Table 1 shows how tones are allocated at the topmost level of the menu structure. Notice that each category at the top level is assigned a specific instrument. As the user progresses down in a category, the instrument remains the same, while the octave and notes vary. This is shown in Table 2. The variability in octaves and notes is as follows: as you go down a menu level you also go down an octave; as you go down a menu item at a given level (except the top level) you go down a note. While moving down a menu level, the cursor is automatically placed at the top item in the new level. This acts as a known home position which should aid in the navigation process.

Table 1: Top Level of Menu Structure

Category	Instrument	Note	Octave
Computer Scientists	Harpsichord	c	4
Physicists	Flute	c	4
Psychologists	Banjo	c	4

Table 2: Level Two of the Structure Under *Physicists*

Menu Item	Instrument	Note	Octave
Atomic	Flute	e	5
Electrical	Flute	f	5
Mathematical	Flute	g	5

2.4 Menu Navigation Methodology

Two groups of subjects are being used in testing the tonally/verbally based menu system. The first group is composed of a number of university students who are normally sighted. Each subject is blindfolded for the duration of the experiment. The second group is composed of a group of non-sighted students at a nearby school for the visually handicapped

All subjects are given supervised training on how to navigate through the menu system using the tool(s) they will be tested with later. At the end of the training period each subject is given a series of randomly selected paths to navigate. The test paths are presented in the form in which they are expected to operate (tonal, verbal or both). The subjects are

[4] XCMDs and XFCNs are special purpose routines written to provide HyperCard and other application programs extra functionality. They are called "resources" in the Apple world and can be easily "pasted" into any application.

then asked to navigate to the designated terminal node in the menu structure.

At any point in the navigation process, the subject may request help in one of two ways through simple mouse clicks. First (s)he can ask for a message from the system which elaborates, using the appropriate audible signals, the exact path to the user's current location. The second help command, again a mouse click, automatically returns the user to the top of the menu structure.

Once the actual tests are begun, all movements are recorded in a file which includes the menu moves made and the number of ticks (measured in 1/60ths of a second[5]) since the last action taken. These data are then saved at the end of the test session for later analysis. In addition to timing and pathway data, we also collect biographical data for analysis by such factors as age and gender.

3 Initial Findings in Sound Based Menu Navigation Experiments

To date, the system has been tested on a small number of sighted users wearing blindfolds. We have not yet been able to schedule the truly non-sighted subjects. Our initial findings are consistent with our expectations. They are briefly:

- Constrained mouse motion and warping are absolutely essential in the navigation of menus by non-sighted users. Even with the use of auditory feedback, movement of the mouse in an unconstrained way leads to general confusion and disorientation by the user. While warping is generally frowned upon by the Apple Interface Guidelines, it is essential in special purpose interfaces.
- Navigation using tonal clues proved difficult for our initial subject population of sighted, blindfolded users. We have yet to collect data from truly non-sighted users, and may see some statistically significant improvement in this domain. We feel that the problems in using tonal navigational clues are based on two main factors. First, our users had no previous experience using such data in the assigned task and therefore had difficulty in associating a tone with an object. Second, the tone generating hardware on the test systems was not of the highest quality. The difference in the same note played at neighboring octaves was sometimes difficult to perceive. It could be that our non-sighted population are more adept at using auditory input in their daily lives and will prove more successful at using tones.
- Speech driven menu navigation was, as expected, the most easily mastered technique. All test subjects were able to navigate or test domain with little trouble with a short training period. It seems clear that given constrained mouse movement and mouse warping, that this technique is quite viable.
- The use of the mouse as a pointing device for visually impaired computer users is not optimal. All users found themselves picking up the mouse and moving it in the upward direction on the desk. Because the cursor was warped to the top menu item in each level, all mouse motion was in a downward direction. Users rapidly found themselves running off the bottom edge of the desk. While users were able to cope with moving the mouse upwards, it seems clear that a trackball would be more useful than a mouse. We feel that this is a very minor change in system hardware and could easily be offered as a no-cost option by the manufacturer.

4 References

1. Blattner, M.M., et al., "Earcons and Icons: Their structure and common design principles," *Human-Computer Interaction*, volume 4, number 1, 1989, pp.: 11-44
2. Borenstein, N.S., (1991) *Programming as if People Mattered*, Princeton University Press, Princeton,NJ.
3. Butler, D. "Describing the Perception of Tonality in Music," *Music Perception*, Spring 1989, 6(3), 219-241.
4. Buxton, W., "Introduction to this special issue on nonspeech audio," *Human-Computer Interaction*, volume 4, number 1, 1989, pp.:1-9.

[5] Naturally, the number of ticks would be recorded in 1/50ths of a second in Europe.

5. Cohen, M. and Ludwig, L. F.,"Multidimensional audio window management," *Int. J. Man-Machine Studies*, 1991, 34, pp. 319-336.
6. Dowling, W. J, and Harwood, D.L. *Music Cognition*, Academic Press, Inc., Orlando, 1986.
7. Edwards, A. D. N. "The Design of Auditory Interfaces for Visually Disabled Users," *Human Factors in Computing Systems (Proceedings of CHI 88)*, ACM SIGCHI, 89-94.
8. Edwards, A.D.N., "Soundtrack: an auditory interface for blind users," *Human-Computer Interaction*, volume 4, number 1, 1989, pp.: 45-66.
9. Gaver, W.W., "The SonixFinder: an interface that uses auditory icons," *Human-Computer Interaction*, colume 4, number 1, 1989, pp.: 67-94.
10. Gould, J., (1988) How to Design Usable Systems, In Helender, M. (Ed) *Handbook of Human Computer Interaction*, Amsterdam: North Holland.
11. Karshmer, A.I., Brawner, P. and Reiswig, G., "An Initial Evaluation of a Sound-Based Hierarchical Menu Navigation System for Visually Handicapped Use of Graphical User Interfaces," to appear in a Springer-Verlag Series on Human Computer Interfaces, 1994.
12. Karshmer, A.I. and Oliver, R.L., "Special Computer Interfaces for the Visually Handicapped: FOB the Manufacturer," the proceedings of EWHCI '93, Moscow, Russia, August, 1993.
13. Karshmer, A.I., Hartley, R.T., Paap, K., Alt, K. & Oliver, R.L., "Using Sound and Sound Spaces to Adapt Graphical Interfaces for Use by the Visually Handicapped," The Proceedings of the 3rd International Conference on Computers and Handicapped Persons, Vienna, July, 1992.
14. Karshmer, A.I., Hartley, R.T. and Paap, K., "SoundStation II: Using Sound & Sound Spaces to Provide High Bandwidth Computer Interfaces to the Visually Handicapped," *SIGCAPH Newsletter*, ACM Press, January, 1992.
15. Karshmer, A.I., Davis, R.D. and Myler, H., "The Architecture of An Inexpensive and Portable Talking-Tactile Terminal to Aid the Visually Handicapped," *Computer Standards and Interfaces*, Vol. No. 5, 1987, North Holland Publishing, pp. 135-151.
16. Karshmer, A.I., Davis, R.D. and Myler, H., "An Inexpensive Talking Tactile Terminal for the Visually Handicapped," *The Journal of Medical Systems*, Vol. 10, No. 3, 1986.
17. Krumhansl, C., & Kessler, E. "Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys," *Psychological Review*, 1982, 334-368.
18. Lerdahl, F. and Jackendoff, R.," A Generative Theory of Tonal Music," *The MIT Press*, Cambridge, Massachussets, 1983.
19. Ludwig, L. F., Pincever, N., and Cohen, M., "Extending the Notion of a Window System to Audio," *IEEE Computer*, August 1990, pp. 66-72.
20. Norman, D., (1986) Cognitive Engineering, In Norman, D., and Draper, S., (Eds), *User Centered System Design*, Hillsdale N.J.: Lawrence Erlbaum Associates, Inc.
21. Ogden, W., (1986) Implications of a Cognitive Model of Database Query: Comparison of a Natural Language, Formal Language, and Direct Manipulation Interface., *SIGCHI Bulletin*, **18**(2).
22. Paap, K. and Roske-Hofstrand, R., "The Optional Number of Menu Options per Panel," *Human Factors*, 1986,28(4), 377-385
23. Paap, K. and Roske-Hofstrand, R., " The Design of Menus," *Handbook of Human-Computer Interaction*, M. Helander, ed. Elsevior Science Publishers B. V. (North Holland), 1988
24. Peterson, I., "The Sound of Data," *Science News*, Volume 127, June 1, 1985, pp.: 348-50.
25. Sagisaka, Yoshinori, "Speech Synthesis from Text," *IEEE Communications Magazine*, January 1990, pp. 35-41.
26. Schmandt, C., Ackerman, M. S., Hindus, D., "Augmenting a Window System with Speech Input\," *IEEE Computer*, August 1990, pp. 50-56.
27. Schmandt, C., McKenna, M.A., "An audio and telephone server for multi-media workstations," *Proceedings of the 2nd IEEE Conference on Computer Workstations*, 7-10 March 1988, Santa Clara, California, pp.: 150-9.
28. Wenzel, E. M., Wightman, F. L., and Foster, H. S., "A Virtual Display System for Conveying Three-Dimensional Acoustic Information," *Proceedings of the Human Factors Society 32nd Annual Meeting*, 1988, pp. 86-90.
29. Wenzel, E. M., Wightman, F. L., and Kistler, D. J.,"Localization with non-individualized virtual acoustic display cues," *CACM*, March 1991, pp. 351-359.
30. York, B.W. and Karshmer, A.I., "Tools to Support Blind Programmers," to appear in *Visual Programming Environments*, edited by E.P. Glinert, IEEE Computer Society Press (1991).
31. York, B.W. and Karshmer, A.I., "An Overview of T^3 - PBE," *SIGCAPH Newsletter*, ACM Press, Number 41, January, 1989.