MULTIMEDIA INTERFACE FOR THE DISABLED (MIND) PROJECT

R J McCrindle and R M Adams
The University of Reading, UK

ABSTRACT

The Multimedia Interface for the Disabled (MIND) project is concerned with developing a set of guidelines and authoring tools, for use by multimedia developers, to enable them to augment their products to encompass the specific needs of sensory impaired users. This paper presents the ethos behind the project and describes the MIND software prototype developed. The MIND prototype maximises the effectiveness of multimedia information delivery, through the provision of an adaptable and easily navigated user interface, which incorporates access to the augmented multimedia information created with the authoring tools.

1. Introduction

Multimedia has been the subject of intense development over recent years. This development coupled with the continually decreasing costs of the enabling technology has resulted in a significant expansion in the use of multimedia. Indeed, it is already noticeable that multimedia applications are impacting significantly upon our environment whether in the workplace, in education or in entertainment (Sloan, 1996).

Multimedia by its very nature of being a multi-modal mix of text, graphics, audio, video, and animation, stimulates our senses of sight and hearing, often in harmony to maximise information delivery. For the majority of the population, a multi-modal mix thus adds value to the information dissemination process. However, there is a small but significant group of the general and computing population who through some form of sensory impairment are unable to exploit fully the capabilities of multimedia.

Research into interface technology for users with various forms of sensory deprivation has revealed a number of specialised techniques, tools and applications that can be used to better facilitate computer interaction for the disabled user. These include Braille screen readers (Alva, 1998), screen magnification tools (Dolphin, 1998), acoustic environments (Lumbreras et al, 1996), speech synthesisers (Dolphin, 1998) and virtual keyboards (Istance, 1996). However, many such developments rely on specifically designed hardware peripherals, which must therefore be 'bolted onto' the software application with which they are to be used. This is due to the fact that the broad spectrum of software products currently in use, and particularly those of a multimedia nature, do not include, interfaces specifically designed to assist the user who has sensory impairment.

Evidence suggests that multimedia applications will become continually intrinsic to our work and everyday lives. Hence there is a very real need to research and define an approach to multimedia product development, such that mainstream multimedia products can be effortlessly used by the full spectrum of the population (Edwards et al, 1995). The MIND project is addressing this need through the provision of mechanisms for selective substitution or augmentation of each media type according to an individual's needs. The approach is being further enhanced through the seamless integration of virtual reality techniques into the interface in order to expand the range of input and output mechanisms by which the user can interact with the computer.

2. Aims of the MIND Project

The MIND project was instigated to further investigate and invite discussion about the perceptions and needs of individuals throughout the entire spectrum of the sensory impaired population, particularly with regard to their use of multimedia applications. The key aims of the MIND project were threefold addressing the authoring, end-user interaction and technological aspects of multimedia development:

Psychological Issue	Interface Considerations
Cognitive Issues	Can established information, theories and methods be applied to the MIND interface? For example, can information processing theory be considered in an effort to conceptualise user behaviour and consequently make predictions about user performance?
Perception & Representation of Information	How information is perceived and represented is of paramount within a system such as MIND. For example, how can the interface enable information projection via particular media types to be accessible to all individuals across the range of disabilities considered?
Attention, Memory & Learning	How can the MIND system directionally focus attention and enable the user to quickly and easily learn and remember the operation of the interface, whatever the disability?
Input Devices	Which input devices are most suitable for a particular disability and how can the interface accommodate these?

Table 1. *Examples of psychological user interface design issues considered*

Sensory Impairment	Possible Physical & Psychological Limitations
Blind	Unable to see the VDU. Complete screen contents may have to be stored In memory for fluent interface navigation. May feel unconfident using the mouse.
Partial Sight	May only be able to see large objects present on the VDU. May be able to only focus on one particular area of the screen at any particular time.
Deaf	Unable to hear sound emitted from PC soundcard.
Hearing Impaired	May be unable to hear certain frequency ranges. May only have monaural hearing with consequent loss of stereo sound perception.
Co-ordination Difficulties	May be unable to utilise the mouse and therefore unable to accurately locate and select controls within a two dimensional screen definition. May be unable to utilise more than a few control keys.

Table 2. *Examples of possible physical and psychological limitations related to impairment*

1. To develop a set of guidelines and authoring tools that can be used by multimedia developers to augment their products to encompass the specific needs of sensory impaired users. These guidelines and tools address navigational and interactive aspects of the interface, as well as provision of compensatory methods for presentation of multimedia information from which sensory impaired users may presently be precluded.

2. To provide the user with an interface that can be adapted personally by themselves to enable navigation,
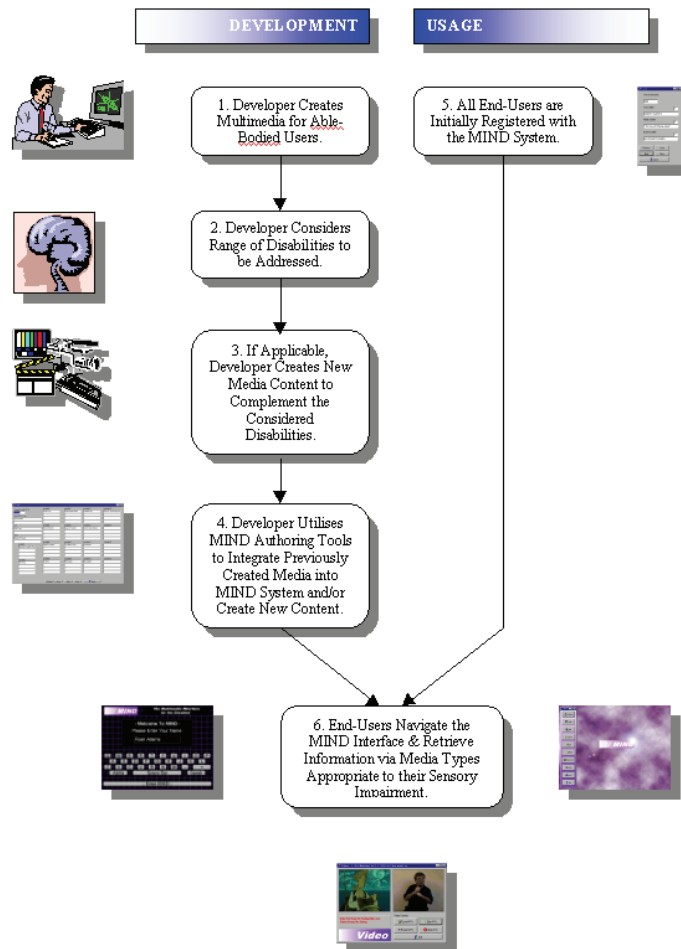

Figure 1: Diagram showing a possible development and usage pathway.

interaction and presentation of previously non-perceivable information in alternative forms according to the specific requirements of their disability.

3. To design and develop a software prototype that addresses the first two aims, whilst operating on any 100% MPC level-2 (Sloane, 1996) compliant personal computer running the Microsoft Windows 95 operating system, to enable maximum uptake and evolution of the system.

Central to meeting these aims was the mapping of the general psychological issues associated with user interface design (see Table 1) to the identified limitations and requirements of the sensory impaired users (see Table 2) within the scope of affordable technology (Eysenck and Keane; Goldstein; McKnight et at, 1993; Preece, 1994).

3. The MIND System Approach

In addition to the design and development of the system architecture, it was deemed important to consider the position of the developers and the end users within a proposed typical development and usage pathway and to account for this with the prototype design. For brevity, this is summarised in Figure 1.

4. The MIND System Prototype

The MIND system is being designed and developed via a combination of rapid evolutionary prototyping and incremental development techniques (Sommerville, 1995; Connel and Shafer, 1995). This approach has enabled a flexible framework to be constructed into which the various interface elements and authoring tools can be progressively integrated. Such an approach has also enabled user feedback to be incorporated into the system features from an early stage. To facilitate rapid and flexible development, within a visual integrated development environment that offers customisable multimedia components, Borland's Delphi version 3.0 (Borland, 1997; Osier et al, 1997) was chosen as the development language. To enable maximum uptake the MIND system has been designed to operate on any 100% MPC Level 2 compliant personal computer running the Microsoft Windows 95 operating system. The key features currently incorporated into the prototype are summarised in the following subsections and Figures 2 to 11.

4.1 LOGGING ONTO THE SYSTEM

The 'Log-in' form provide the means of entering the MIND system. It consists of a virtual keyboard and a data entry box. A toggle-audicon enables the Querty arrangement (see Figure 2) of the keyboard to be changed to an alphabetical format (see Figure 3) and vice versa. Each time a letter of the alphabet, represented by an audicon, is selected, the chosen character is added to the end of the name string in the data entry box. A blind user can check the text currently entered in the data entry box by selecting the 'Speak' audicon. Navigation of the keyboard may also occur via tabbing from one audicon to the next.
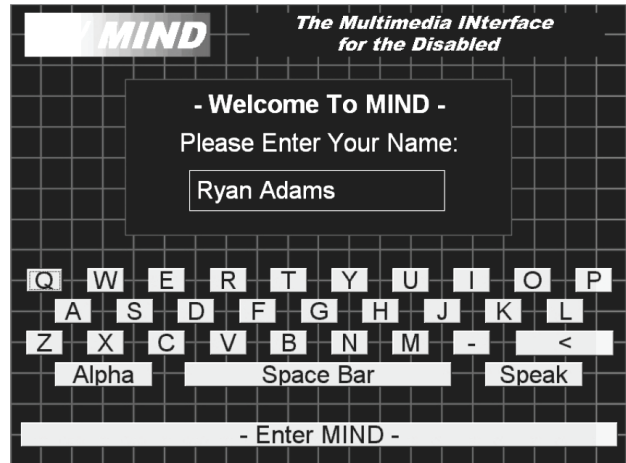

Figure 2: The Log-in Form showing Qwerty arrangement of the virtual keyboard

Alternatively, the MIND system incorporates Microsoft's Speech Recognition engine (1998) thereby enabling the user to enter their log-in name verbally by means of an external microphone. If recognised as a registered user, the system acknowledges the user and automatically loads all associated preferences saved under that user's login name. If a particular name is not recognised then the system asks the user to try again, advising them of the need to speak as clearly as possible.
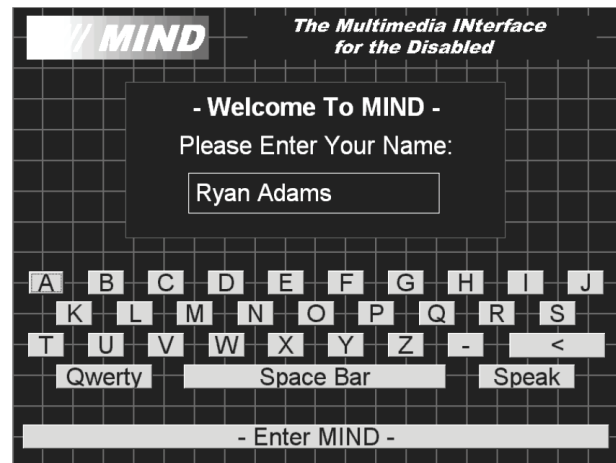

Figure 3: The Log-in Form showing Alphabetical arrangement of the virtual keyboard

The preferences of each user are stored in a 'User Registration File' via the 'User Registration Tool'. This tool allows the addition and editing of users currently registered with the system. It also compiles users' names into the special 'grammar file' needed by the speech recognition engine to determine which log-in names are valid.

4.2 THE MAIN MENU AND INTERFACE

The multimedia interface consists of a resizable and adaptable 'Main Menu' window or form containing the buttons from which all facilities that provide access to the multimedia components can be selected (see Figure 4). This form is completely resizable such that the buttons also resize, thereby maintaining their spatial relationship along the vertical axis. The font describing the buttons' functionality also resizes in order to provide partially sighted users with the largest typeface possible within the vertical screen definition. Resizing the main menu to cover the full width of the screen also aids in the spatial location of buttons for blind users since they only have to consider the vertical axis in order to locate a particular button with the mouse cursor.
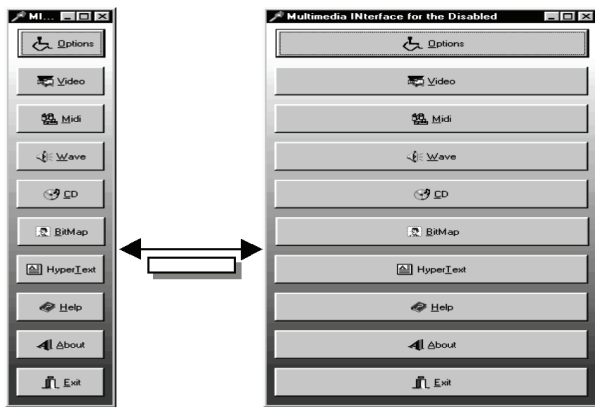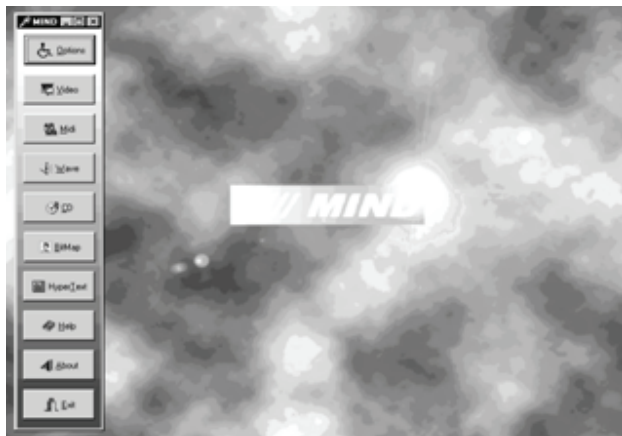

Figure 4: The Resizable Main Menu Form

Figure 5: *The Underlying Backdrop with Windows 95 Taskbar Removed*

To prevent accidental selection of applications outside of the MIND system by blind users, the Windows 95 Taskbar is automatically removed from view while the MIND system is running. The Main Menu Form is also placed upon a backdrop (see Figure 5) so as to eliminate the possibility of blind users selecting Windows 95 desktop items with the mouse cursor.

4.3 SELECTION OF FEATURES

All features may be selected in the normal manner by positioning the mouse cursor over the corresponding screen button and depressing the left mouse button. However, the buttons are also implemented as 'audicons', which are entities that represent a certain functionality by auditory as well as visual means. As the mouse cursor is moved over a button a digital sample of a voice describing that particular button's operation is emitted. Alternatively, the options available within the main menu or tools may be located by means of 'tabbed' keyboard input coupled with explanatory auditory information.

This simple keyboard interaction has also been designed to assist users with co-ordination impairment since it eliminates the needs for physical dexterity and accuracy often associated with the use of a mouse. Additionally, blind users who do not feel completely confident with the mouse may also prefer this method of interaction. Future developments of the system may also consider 'headset' or 'eye stare' selection of features on a simplified virtual keypad.

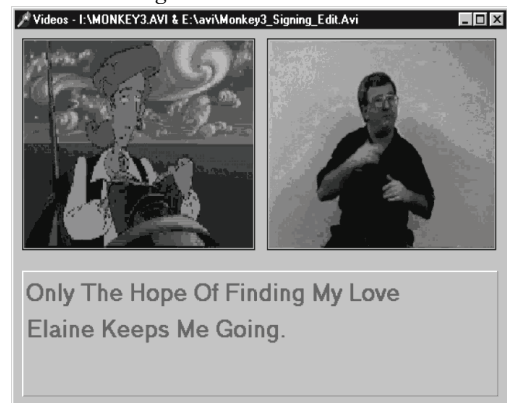*Figure 6: The Video Form*

Figure 7: Video Form with Enlarged Subtitle Area

4.4 NAVIGATION WITHIN MIND

Many of the navigational aspects of the MIND system have been determined by the needs of visually impaired and blind users. The MIND system follows a simple 2-tiered hierarchical design. This means that in order to navigate

from any one form or multimedia representation to another, the user must return to the main menu. Additionally, when the user selects an option from the main menu the resultant form or window is displayed 'modally'. This means that the current form must be closed before the user can reactivate the main menu. Although, this may appear to be somewhat tedious, it ensures that the user is prevented from becoming 'lost' within the system.

The selection of, or exit from, any of the forms available within the main menu is also accompanied by auditory information describing the users actions and the current status of the system. Additionally, to prevent accidental selection of applications outside of the MIND system, the Windows 95 Taskbar is automatically removed from view whilst the MIND system is running.

4.5 CONSISTENCY AND DEFAULTS

Figure 8: The Subtitle Authoring Tool

To provide a consistent interface, by default all the attributes pertaining to the buttons on the 'Main Menu' form are extended to every button throughout the MIND system. However, the 'Options Menu' that is currently being implemented, will provide methods whereby the attributes of each button within the interface can be adapted to suit the preferences of individual users.

4.6 THE VIDEO FORM

One of the key multimedia features implemented to date is the 'Video Form'. This form enables the synchronous display of two videos and a corresponding subtitle file (see Figure 6). The left video constitutes the

'Main' video, whilst that to the right may be an accompanying 'Signing' video, thereby translating the audio within the main video to sign language for comprehension by deaf users. Additionally, a subtitle file may be created using the Subtitle Authoring Tool and then displayed in parallel with the main video. The subtitles may also be displayed in different colours to indicate which individual within the main video is currently speaking or in order to depict background sound. The video form and its components are completely resizable to aid partially sighted persons, and if required the subtitling may be expanded to fill the entire lower half of the form (see Figure 7).

At present, the video form only supports media in AVI format but future developments will allow the presentation of MPEG encoded video. For demonstrative purposes, a suite of media was assembled. The main video was a short promotional video for a computer game. The second video was a British sign language (BSL) translation of the main video which was also complemented by an associated subtitle file.

The 'Subtitle Authoring Tool' (see Figure 8) provides an interface whereby subtitle text can be easily entered along with a time of activation and a particular colour. The use of colour within the subtitles helps identify the different characters speaking within a video and may be defined by assigning colours to an integer value which is then specified using the tool. For example the value '1' corresponds to the Colour 'Black' for the subtitles defined in Figure 8. The developer can browse through the list of subtitle records by selecting the 'next' and 'previous' buttons.

4.7 THE CD FORM

The Audio CD Database Tool has also been successfully implemented. This enables the provision of a CD form and associated interface from which audio CD tracks can be played in parallel with a corresponding video and subtitle file (see Figures 9 and 10). The CD form consists of: a main resizable window, a display area for subtitles, an area for AVI presentation and an area where current CD information can be displayed.

Figure 9: The CD Form with an
Associated Video and Subtitle File

*Figure 10: The CD Form without an
Associated Video and Subtitle File.*

All CDs that are registered with the MIND system by use of the CD Database Tool (see Figure 11) will automatically be recognised and relevant files loaded. The information that can be stored about each CD is currently:

• The title of the CD and a corresponding auditory description.

• The artist of the CD and a corresponding auditory description.

• The label of the publisher who released the CD and a corresponding auditory description.

Each individual CD audio track may also be associated with:

• Textual and auditory information describing the title of the track.

• A MIND subtitle file that contains textual information describing the contents of the track, or which is a 'dictation' of the voiced words within the CD track.

• A Video file, which may contain relevant information such as a signing translation or a music video.

The 'Speak Info' audicon informs a blind user of the current track playing. If a WAV file has not been associated with the currently selected track then the system will read out the text by utilising the Text-to-Speech (TTS) engine. The MIND system utilises both Microsoft's Speech API (1998) and a set of shareware components collectively called DTalk (1998; Barbosa, 1998). DTalk consists of a Text-to-Speech component to enable talking applications and Grammar/Speech Recognition (SR) components to make applications listen and understand. The DTalk components interface to the Microsoft TTS and SR engines present in their Software Development Kit (SDK) (1998).

Table 3. *Features under development and incorporation within the MIND prototype*

Feature	Tools being Considered
The MIDI Form	A designated video area where information about the current MIDI tune playing can be displayed (e.g. display of a composer purveying the current tempo and auditory experience). A visual piano keyboard, whose keys change colour, indicating the current notes in play. A scrolling musical score, which highlights the notes currently in play. Textual information describing the style, tempo and key of the music and who it was comp[...] by.
The WAVE form	A designated subtitle area where dictation of spoken word within the sound sample (if any) or a description of the actual sound can be displayed. Graphic display of the waveform output. Animated graphic equalisers showing intensity for various wavelength intervals at any specific time.
The BITMAP form	Each bitmap file is to be associated with auditory information describing the picture. The facility to enlarge an image will be provided to aid those persons with partial sight.
The HYPERTEXT form	A form which allows HTML files to be loaded and displayed as well as providing methods by which disabled users may interact and direct the information flow. Speech outputs to be achieved by utilising Microsoft's text to speech engine.
The OPTIONS form	A form that enables the interface and its facilities to be further adapted to suit the user, for example by allowing default file locations to be specified or volume levels for specific devices to be altered.
OTHER tools	Using '*stereo sound*' to provide a guidance mechanism to the spatial location of audicons. Providing '*hot keys*' to enable the user to be transported to a specific location within the interface. The use of Microsoft's *Speech Recognition Engine* to eliminate the use of mouse or keyboard input.
VR tools	The use of **ambiosonic sound** to generate an enhanced 3-D navigation system, incorporating variable volume levels, echoing, and narrative descriptions of the commands selected. Development of an **eye stare virtual keyboard** incorporating the limited key commands required for navigating the system. Incorporation of **speaker vibration** as an output mechanism. Extension of a 2-D multimedia interface into a **3-D virtual world**.

5. Work in Progress

Work within the MIND project is ongoing. The features currently being developed and incrementally incorporated into the MIND system are summarised in Table 3.

6. Testing

Testing and user evaluation is essential for a system that relies extensively on the success of its interface. With this in mind feedback has been obtained from a profoundly deaf subject whose job role is to use video and signing communication strategies to help individuals who have recently lost their sense of hearing. Feedback has been extremely positive and the person in question is interested in using the software within their line of work. Other visually and aurally impaired users have also been consulted during the project. However, it is acknowledged that to date, due to the constraints of the project, testing of the system has been of a qualitative rather than quantitative nature. More extensive and formal evaluations of the MIND system, with planned and will be conducted within the next couple of months. This form of testing, with large numbers of

individuals and controlled experiments will provide more statistically significant results. However, in the meantime all qualitative feedback obtained will continue to impact on the development of future prototype versions.

7. Summary

This paper has briefly introduced the work associated with the MIND project. Substantial progress has been made in a very short space of time and implementation is ongoing. A demonstrable prototype, successfully incorporating the features described in Section 4 is available, and is stimulating discussion and feedback into the needs and expectation of users with sensory impairments. It is expected that the results and ideas stemming from this project, which is being conducted with very limited resources, will act as a pump-priming activity for a more detailed and funded study into ways to adapt multimedia software products to meet the needs of sensory impaired users. The sensory impairments considered within the

Table 4. *Summary of sensory impairments, possible limitations and compensatory methods*

Sensory Impairment	Possible Physical & Psychological Limitations	Compensatory Computing Methods Implemented Within MIND
Blind	Unable to see the VDU. Complete screen contents may have to be stored in memory for fluent interface navigation. May feel unconfident using the mouse.	The provision of audicons, sound guidance mechanisms, TTS & SR. Removal of Windows 95 taskbar. The provision of a 2-tiered hierarchical design. The 'modal' presentation of forms focusing attention. Keeping all form contents to the bare minimum without compromising on functionality. The provision of audicons that cover the whole vertical screen definition. The facility to utilise a small portion of the keyboard as an input device.
Partial Sight	May be only able to see large objects present on the VDU. May be only able to focus on one particular area of the screen at any particular time.	The ability to scale the screen contents. Audicons and controls are located in close proximity to one another in 'clusters'. Forms are represented and displayed 'modally', eliminating the need to see the whole interface and subordinate child forms at the same time.
Deaf	Unable to hear sound emitted from PC soundcard.	Provision of compensatory methods such as subtitle presentation.
Hearing Impaired	May be unable to hear certain frequency ranges. May only have monaural hearing with consequent loss of stereo sound perception.	No solution at present. Real-time processing of digital sound samples could be performed to evade the frequency range the sensory impaired individual cannot perceive. No solution at present. Real time processing could convert stereo sound to mono. More simply, the line-out wires from the soundcard could be fused to produce mono sound.
Co-ordination Difficulties	May be unable to utilise the mouse. Therefore such individuals may be unable to accurately locate and select controls within a two dimensional screen definition. May be unable to utilise more than a few control keys.	The provision of audicons. The ability to scale audicons to cover the complete horizontal and vertical screen definition. The ability to use the keyboard as an input device. The MIND system only requires two control keys; a navigation and select key. Speech Recognition.

MIND project to date, the possible physical & psychological limitations of individuals possessing these impairments and the compensatory methods provided by the MIND system prototype are summarised in Table 4. Work is continuing in the areas described in section 6. There is also scope in future studies for using virtual reality techniques to extend the MIND system to address the specific needs of the elderly and of individuals with learning difficulties.

REFERENCES

A Sloane (1996), Multimedia Communication, McGraw Hill, London.

Alva Technologies Netherlands (1998), Home Page, *http://www.alva-bv.nl/.*

C Barbosa (1998), DWave Delphi Component, *http://www.digiserve.com/cmab/.*

Borland International (1997), *Borland Delphi 3 Object Pascal Language Guide*, Borland International.

J Connel, L Shafer (1995), *Object-Oriented Prototyping*, Yourdon Press.

Delphi Super Page (1998), Components For Delphi 1, 2 & 3, *http://www.delphideli.com.*

Dolphin Computer Access (1998), Home Page, *http://ww w.dolphinaccess.com*

Dtalk, Speech Components For Delphi, *http://www.o2a.com/*

W K Edwards, E D.Mynatt, K Stockton (1995), Access to Graphical Interfaces for Blind Users, *Interactions,* pp. 54-67, January.

M W Eysenck, M T Keane, *Cognitive Psychology*, Psychology Press (3rd Edition).

E B Goldstein, *Sensation & Perception*, Brooks/Cole Publishing (3rd Edition).

H O Istance, C Spinner, P A Howarth (1996), Providing Motor Impaired Users with Access to Standard Graphical User Interface (GUI) Software via Eye-Based Interaction, *Proceedings 1st European Conference on Disability, Virtual Reality & Associated Technologies*, pp. 109-116.

S Jarol, D Haygood, C D Coppola (1996), *Delphi 2 Multimedia Adventure Set*, The Coriolis Group.

M Lumbreras, M Barcia, J Sanchez (1996), A 3D Sound Hypermedial System for the Blind, *Proceedings 1st European Conference on Disability, Virtual Reality & Associated Technologies*, pp 187-191.

C McKnight, A Dillon, J Richardson (1993), *Hypertext: A Psychological Perspective*, Ellis Horwood.
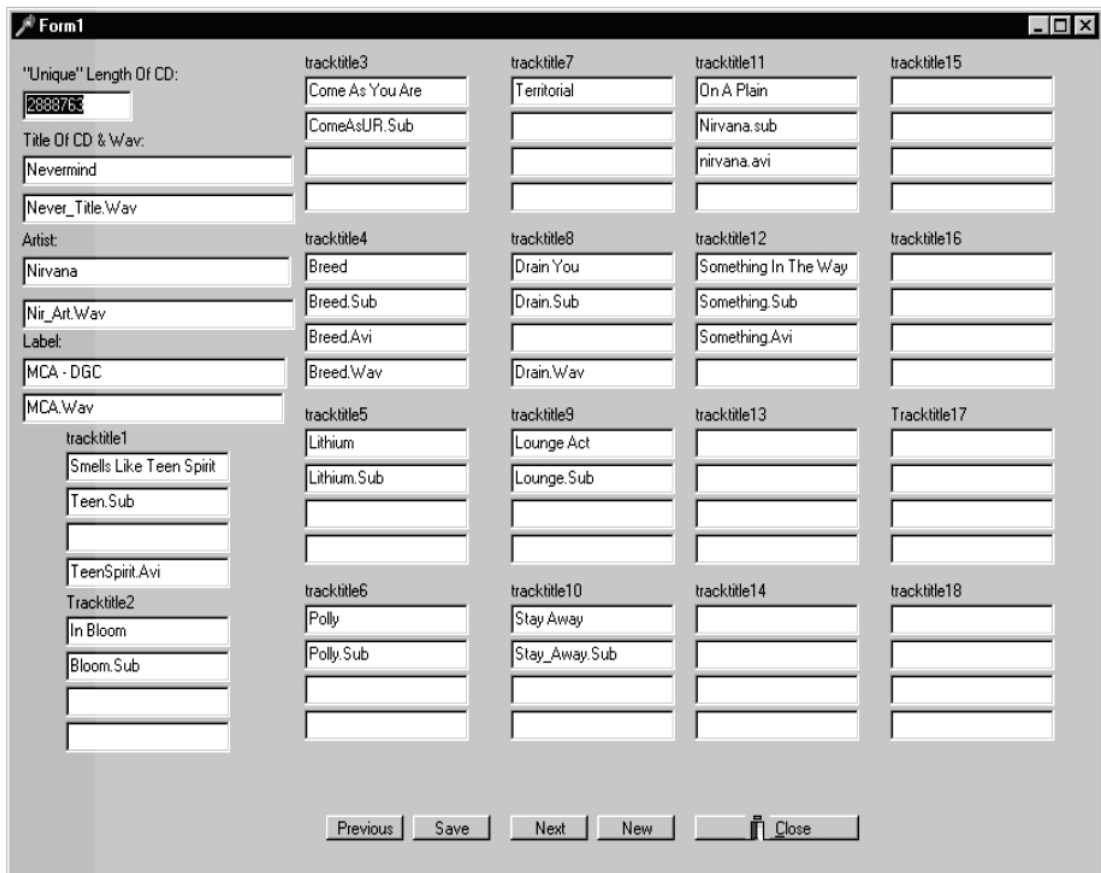
Microsoft, Speech Development Kit, *http://research.microsoft.com/research/srg/*

Figure 11: The Audio CD Database Tool

MIDAS, Sound System For Windows, *http://hornet.org/*

L A Mothe (1993), *Games Programming In 21 Days*, Sams Publishing.

D Osier, S Grobman, S Batson (1997), *Teach Yourself Delphi 3 In 14 Days*, SAMS Publishing.

Jenny Preece et al, *Human Computer Interaction*, Addison-Wesley (1994)

M Rothberg, T Wlodkowski(1998), Multimedia: Making it Accessible to Blind Users, *http://www.dinf.org/csun_98 /csun98_04.htm*

I Sommerville (1995), *Software Engineering*, Addison-Wesley (5th Edition).

D M Tanner (1998), Designing Accessible Computer Systems for the Blind Without Expensive Adaptive Hardware, *http://www.dinf.org/csubn_98/csun98_109.htm/*

BIOGRAPHIES

Rachel McCrindle (photo on CDROM) is a lecturer in Computer Science at the University of Reading. She has a BSc in Botany (Plant Physiology) from Liverpool University, a MSc in Information Technology for Manufacture from Warwick University and a PhD in Computer Science from Durham University. She lectures and researches in software engineering, social impact of information technology, graphical user interface design and multimedia technology.

Contact information:
    R. J. McCrindle
    Department of Computer Science
    The University of Reading
    Whiteknights, Reading RG6 6AY, UK
    Email: R.J.McCrindle@reading.ac.uk

Ryan Adams (photo on CDROM) has a BSc in Computer Science and Psychology from the University of Reading and is currently working for the University, in association with Industry as a Teaching Company Associate, researching and developing MPEG-5 applications.

Contact information:
    Ryan Adams
    Department of Computer Science
    The University of Reading
    Whiteknights, Reading RG6 6AY, UK
    Email: R.M.Adams@reading.ac.uk