Substitution for a Restricted Visual Channel in Multimodal Computer- Human Dialogue

DAVID R. HILL, MEMBER, IEEE, AND CHRISTIANE GRIEB

Abstract — **This paper describes a workstation that uses a multimodal approach to screen substitution for the visually impaired computer user—programmer or applications specialist, and presents the results of an experiment that compared performance using a working prototype of the new design to performance using a conventional key -controlled talking terminal. An improvement in performance that was both statistically and practically significant was found. Furthermore, the new device was preferred by members of the target population. The study also noted that a computer screen is a less than optimal substitute for printed documents, especially books, even for sighted users.**

I. INTRODUCTION

INCREASING numbers of interesting and remunerative careers depend, these days, on an ability to use computers. Jobs are available for those trained in computer skills, and also for those with skills in other fields who are able to cope with the computer aids that are available. Companies find that appropriate applications of computers generally increase productivity and profits while allowing them to provide a better service. Many new products and services arise from the direct needs for new computer hardware and software to provide appropriate aids in manufacturing, professional activity, banking, and the like. One result of this pressure has been to find better ways of making computer power available to those who need it, either as computer professionals generating new products, or as end users applying leverage to their brain power by exploiting these products in their businesses. This is the fundamental reason for the blossoming of computer-human interaction as a subdiscipline, with roots in computer science, psychology, linguistics and electrical engineering. Innovation in computer-human interaction (CHI) can make computers more accessible, productive, reliable and satisfying to programmers and end-users alike.

Manuscript received February 17, 1987; revised November 3, 1987. This work supported in part by Esso Resources Ltd., in part by the University of Calgary Special Projects Fund, and in part by the Natural Sciences and Engineering Research Council of Canada under grant A5261. D. R, Hill is with the Man-Machine Systems Laboratory, Department of Computer Science, University of Calgary, Calgary, AB, Canada T2N 1N4, C. Grieb was with the Man-Machine Systems Laboratory. She is currently with IBM Deutschland GmbH, Entwicklung und Forschung, Schönaicher Strasse 220, 7030 Böblingen, West Germany.
IEEE Log Number 8818923.

In this context, careers based on computers seem increasingly attractive to people suffering from visual impairment. The use of magnified fonts, Braille printers and speech synthesisers, coupled with the ability to touch type and call for a little help, have allowed such people to be competitive in ordinary careers in computing. Special talking terminals were developed to allow line and even simple screen editors to be used reasonably effectively based on the use of keys to provide incremental control of the system cursor ("step" keys); various function keys to control the facilities offered, including the form of speech feedback (spelled or spoken, on or off, etc.); and speech itself to act as a substitute for the normal visual feedback needed to find and correct material being edited, or read documentation, help text, and the like.

There are some rather subtle aspects to this kind of operation, especially given the substitution of audition for vision, that may not be immediately obvious. For example, more than one conceptual cursor is involved. Apart from the obvious system cursor, which marks the place where changes will occur and may determine (for a talking terminal) what will be spoken next, there is a user cursor that corresponds to the point of regard for the sighted user, and may be used in a variety of ways. For example, it is implicitly used in visual search to determine where next to move the system cursor, or as the basis for checking context when deciding whether or not the system cursor is correctly positioned. It is also used in simple reading, while the system cursor keeps track of what may be called the "working location." Other cursors may be associated With the extent of a region that is to be identified for some purpose, or as place keepers. In reading a real book, information is available in the form of book marks, and location within the book (how much of the book lies either side of the pages currently visible) or within the page. This information is valuable and usable when trying to find something in a large body of text. Sighted readers often have a clear visual memory of where to find something. Such information is not readily available even to sighted readers, once they access text using a computer because page layout, for example, may vary during scrolling. Scroll bars may give an idea of whereabouts the reader is within the overall text file, but usually only in relative form. Deciding on the absolute size of the file, getting a feel for

0018-9472/88/0300-0285$01.00 ©1988 IEEE

how far to advance to reach some desired point, may be difficult and very likely will depend on counting pages or using a key term or heading, which is only part of how people read real books. This sort of limitation explains why many people, even experienced computer users, prefer to print out a hard copy of computer text for proof reading, debugging, or reference. Benest and Jones [2], in their work on the computer emulation of books, have given a good account of some of the attributes people use in reading books. The topic is also considered in [7] and [5, ch. 3], where problems of existing computer substitutes based on both screen and talking terminal access are discussed.

When text is presented by computer, for reading, as part of an edit session, and so on, many of the cues that facilitate such activities as browsing, searching, and verification during other kinds of text access (for example, in books or files) have been changed or even eliminated. Thus even normal readers labor under some constraints when accessing or editing text using a computer, as noted. On the whole, users seem to feel that the benefits outweigh the disadvantages. They like the ease with which corrections may be made, and the convenience of various other facilities including fast access and text manipulation This has clearly been true, to a lesser extent, for the visually impaired user, even though such a user works at a considerable disadvantage compared to the sighted user. Thus, when reading a (Braille) book or other document, tactile cues allow a blind reader to judge how big it is, whereabouts the current location is relative to the start and end of the text, and so on. When reading computer text, the sighted reader can use visual feedback (line counts and numbers, for example) to substitute for the missing cues, but the blind user only obtains such information by laborious means.

The advent of new forms of interaction such as direct manipulation, which place a vastly increased emphasis on much more demanding forms of visual feedback, threaten to make the relative disadvantages suffered by the visually impaired quite overwhelming. At least with step keys and search commands, the visually impaired user can count, touch-type, and receive relevant auditory feedback; although tasks like manipulating more than one cursor, or ensuring that the intended occurrence of a word has been found are more problematical. Equally, learning all the function key combinations and their effects (especially if layouts or assignments are awkward) presents problems. One commercially available talking terminal had key combinations that were physically difficult to span with a normal hand. A sighted user can easily monitor actions by direct visual feedback and has fewer commands and functions to worry about anyway. But if simple access can be difficult, how does someone who cannot see the screen use a "mouse" to position a cursor or select an option? How does such a user interpret the arbitrarily ordered string of characters and control codes returned from a host during optimized screen update? How can that user keep in mind the different functions and purposes of material in multiwindow operation, if it can only be accessed serially, in the order of the character raster or host input? And how distracting will it be to receive continual time updates and messages confirming checkpoints, over the audio channel, when concentrating on composition. The alternative may be worse. If the audio output is disabled for incoming messages, important error or system information ("System down in 2 min, please save files and log off."), may pass unnoticed.

The Touch 'n Talk workstation, being developed on the basis of research carried out at the University of Calgary, provides some of the advantages and mechanisms of direct manipulation access and control for visually impaired computer users (programmers or end-users), by substituting touch, synthesised speech, and proprioceptive feedback for vision. Our experiments suggest that visually impaired users perform editing tasks better, can access menus and the like more effectively, and are more satisfied by such a device than when using a conventional key-based talking terminal.

II. SYSTEM OVERVIEW AND RATIONALE

A. Introduction

The Touch 'n Talk I workstation, in its current form is based on an Atari 1040ST personal computer. To provide the special forms of input and output needed to carry out its functions, a talker, a textured pad, and a device to send and receive data over a telephone line are also all attached to the computer. Options include a Versabraille[1] machine, a printer, and a Braille printer. Fig. 1 shows a photograph of the textured pad currently in use. The pad, which may vary in design, is central to the system, providing a tactile/proprioceptive model of the screen that, coupled with text-to-speech translation, becomes a pseudodisplay. Although a prototype system has been implemented and tested (see Section III), what follows is a brief user-oriented view of a full system currently in an advanced state of development in the first author's laboratory. Not all parts are fully implemented, but enough has been tested to form a standalone editing workstation that should be fully operational by the time this paper appears in print.

B. Speech Output

The talker (a resonance analogue speech synthesizer, together with driving programs in the system), is able to convert any written material into equivalent spoken output, at various rates of speech. Variations in rhythm, intonation and speaker characteristics are possible.

C. The Pseudo-Display

The current textured pad (Fig. 1), shown diagrammatically in Fig. 2, is 12 in square and has a central portion about 11.5 in by 9 in laid out with 44 horizontal grooves

[1] Versabraille is a registered trademark of Telesensory Systems Inc.

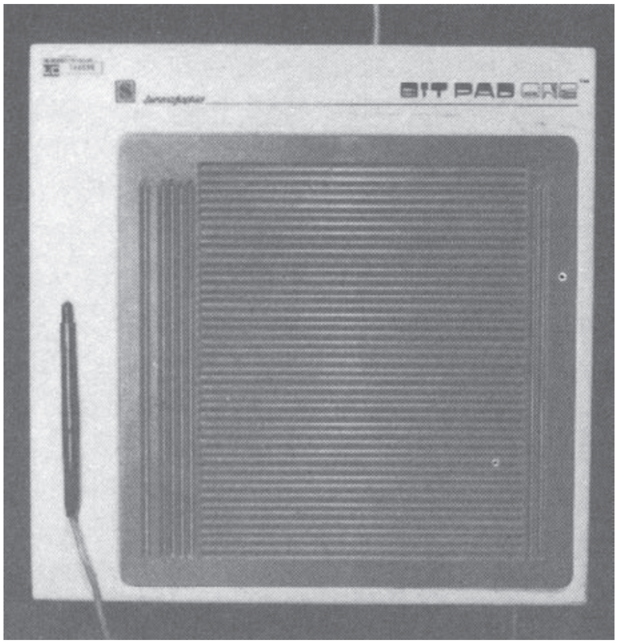

Fig. 1. Current textured pad

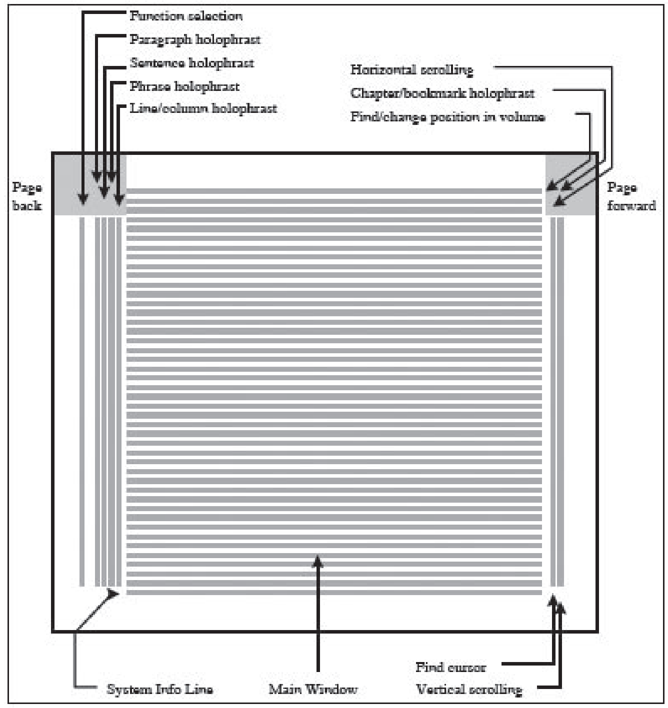

Fig. 2. Diagrammatioc representation of the textured pad

(rows) in a pattern analogous to the lines of text or Braille in a book. These are flanked by vertical grooves (columns), five on the left and two on the right. No specific fixed characters are represented in the grooves, or anywhere else on the pad. However, certain regular patterns may be felt to help in keeping track of position and movement when feeling the pad. The pad provides a physical reference for what may be thought of as a virtual screen on which text and function selections may be "displayed" and/or activated. The pad, backed by synthesized speech, thus forms a pseudodisplay. The real screen (CRT display) is also made to show what is placed on the virtual screen, partly to help in software development for the system by sighted programmers, and partly to allow sighted colleagues to collaborate with visually impaired users. The particular material referenced at any given time is determined by the computer and user working together. The computer is able to detect gestures (rates and kinds of movement) as well as position, so that a variety of flexible, convenient forms of control and pseudodisplay are possible. Thus page turning may be associated with the action of stroking a top corner of the pad, or the action specified in a column may vary depending on just how the column was entered. Of the 44 rows on the tablet, the lower 41 represent a "window" that reveals a portion of a file or document, or provides a working area for other purposes. The characters that are revealed in the window are placed in simple one-to-one correspondence with distinct physical co-ordinates on the textured pad. At any given time, a particular character within the material being dealt with is remembered by the computer as its working location. A special flag, called the system cursor, marks this position so that the user can also find it. The user's working location, corresponding to the "point of regard" or eye fixation point for normally sighted users, can also be marked by a special flag, the user cursor, so that the system can keep track of it. The system and user cursors may or may not coincide, depending on what operations are taking place. It is always possible to move the system cursor to the user cursor and vice versa, or to find the system cursor, or to remember the current position of the user cursor and restore it to its remembered position later on. In addition to setting up a main working area, provision is made for natural convenient access, function selection, and control. Two columns run from the top edge of the pad to the bottom, on the right-hand side. One of these allows the system cursor to be found or manipulated quickly. The other allows fine control for the position of the window within a file or document, if files or documents are being viewed. Five more columns run down the left-hand side of the pad. Four of these provide a structured view of whatever text material is being examined or edited and are called holophrast columns. A holophrastic display presents material in condensed form, as nodes, and provides a facility for expanding the nodes. Such a structure may be hierarchical. The four holophrast columns of the Touch 'n Talk operate on single level nodes and allow: headings/paragraphs; sentences; phrases; and lines or column entries to be detected and also spoken, if desired, without having to touch the main text area (working area). The fifth column, on the extreme left of the pad, provides a variety of soft function selections, including asking for help and calling different windows onto the pad. In addition, page turning is accomplished by stroking the top left corner of the pad to page back, or the top right corner to page forwards, as if picking up the corner of a real page. This requires deliberate action, and avoids accidental page changes, acting as a gestural analog of page turning. Other functions also depend on the recognition of such gestures, rather than explicit, nonredundant function selection using simple buttons, partly to make selection robust and partly to tap into the normal habits of thought possessed by users. Within the main working area, what is spoken depends on the movement of the user cursor which is controlled by the user's finger moving along the grooves accompanied by the tip of a digitizing stylus. The use of a stylus originally seemed a regrettable compromise with less than perfect technology, but experiment suggests that its use is actually preferrable, as explained in the discussion of the experiment below. Nevertheless, "finger position" will be used in talking about user cursor control, since this is the subjectively important aspect. The tactile cues are important.

Thus finger position determines what is spoken, and its rate of movement determines the rate of speech. Above a certain rate it is assumed that the reader is skimming, and only those important words that can fit into the time available are spoken. If the user moves very slowly, it is assumed that a spelled version of the current word is needed, and the system proceeds in spelled speech mode. If the user stops, the speech stops. If the user then taps at the same location, the most recent (current) word is repeated as it was last spoken (spelled or normal). If the user taps twice in quick succession (a double tap), then the system cursor is made to coincide with the current user cursor and its old position is remembered by placing a third special flag called the mark there. If the user moves slowly back along the reverse direction of reading, the words are repeated in reverse order (presumably some intended point has been passed). But if the user moves back quickly, it is assumed that this is a "move to" gesture, and no further speech or action is generated until the user cursor reaches an active function, an active holophrast node, or resumes normal movement within the text area. A similar distinction is made for movements in other directions.

At the right, the column nearest the text area allows the system cursor to be found. Touching within this column causes a tone to be generated. In moving up and down the column, the frequency of the tone varies, decreasing as the user cursor gets closer to the line position of the system cursor, and becoming zero or null (no tone) when the user cursor is opposite the text cursor line. Moving left into that line of the text area then repeats the process of tone generation with the null occurring at the character marked by the system cursor. In normal mode, the cursor will be on the first character of a group of characters surrounded by white space (tabs, spaces, end-of-line characters, etc.)—a word. In spelled-speech mode, the cursor will be positioned on the character last accessed by the user cursor. If the user happens to touch exactly on the null point during such a search, a slight movement either way will quickly reveal this. If the system cursor is not within the current working area, the top or bottom of the line-finding column will be reached before the tone vanishes. Tapping twice in rapid succession will reposition the window so that the system cursor line is in the middle of the window (for nonpaged material), or the window will be moved to the page in which the system cursor appears. In either case, the cursor is then readily found, but definite action is required to move off the current page/window. A system message is generated that keeps a note of the "distance" moved, and may be accessed in the system message window.

A double tap at the location of the system cursor will exchange the mark with the system cursor. The swapping process may be repeated indefinitely, but provides a means of locating the mark, if this should prove necessary. With the user cursor at the system cursor, the user simply swaps the system cursor with the mark and then finds the system cursor again.

The second column at the right is used for fine vertical positioning of the window within a body of text. Stroking up or down in the column and then tapping once (to signal readiness) moves the window by an amount physically equivalent to the finger movement. If paged material is being dealt with an audible beep sounds when the top or bottom of the window reaches the top or bottom of the page, as appropriate, and the window stops. If a further stroking movement is made in the same direction, movement continues again, and a new beep signifies if and when the middle of the window reaches the top or bottom of the page. At this point, the window cannot be moved any further using the vertical positioning mechanism. This is useful in maintaining consistency in layout while still allowing material crossing page boundaries to be dealt with conveniently. To move to a previous or subsequent page, the page turning mechanism must be used. If the page is turned, the window will then be positioned with its top edge coincident with the top edge of the new page.

For unpaged material the rightmost column simply allows continuous scrolling. However, based on our experience, we believe that all computer text should be in paged form, and unpaged material should be "published" (by passing it through at least a simple text formatter) before being made available for viewing.

The holophrast columns allow a convenient structured view of, and access to, material in the working area. Running up or down one of the holophrast columns produces no effect until the user cursor is opposite a line containing the start of at least one of the specified elements (heading/paragraph, sentence, phrase, or line/column entry). This position is a holophrast node for the kind of unit in question. A beep is produced when a node is encountered the first time following some other activity. Tapping at the node after it has been identified causes the first relevant unit on the line to be spoken until its end, or until the user's finger is raised. In the latter case, touching the pad surface again resumes speech output at the point of interruption, just as it would during normal text access in the main area. The speech rate during holophrastic access is arbitrarily fixed, of course, since the user cursor (finger) is not moving. On reaching the end of the unit, a beep is generated if there is a further unit starting on the same line, as a signal to inform the user of this fact. Raising the finger and lowering it again, at this time, will cause the next unit to be spoken, a process that repeats if further units also follow. If the spoken unit runs over to the next or subsequent lines, the beginning of the next unit of the same type is found by continuing on down the column. Moving up or down the column during output terminates output and allows the user to find the next line, in the direction of search, that contains the start of another unit. Tapping the current holophrast node will repeat the current unit from the beginning, while tapping twice in rapid succession will force the speech output to begin again at the start of the first unit on the line. As will be noticed, gestural control is again being used. The holophrast method provides convenient access to the rows of columnar data, as well as access to structured text and program text. If, on the other hand, it is necessary to access successive items in column order, the user can find the intended column in the main window area, and then scan the column items vertically to produce the required output. The system is able to recognize most columnar data because of its regular layout and extra spacing, so that column entries may be handled properly, by row or by column.

Finally, the three top horizontal grooves in the working area are also special. If the text being accessed is regarded as a book, the position within the book may quickly be detected, marked and/or changed by using the top groove. The length of the groove may be though of as representing the largest possible book—900 pages (material so large that it exceeds this notional size should be broken into numbered volumes for separate access). A volume being viewed is represented by an appropriate proportion of the groove, starting at the left end. Within this portion a location tone, similar to the find-cursor tone, allows the user to find where the book is open. Beyond the end of the volume, the tone drops to a constant low tone. To place a bookmark at the current location, the user simply taps once in the vicinity of the point located. The system automatically labels the bookmarks, and enters them in the holophrast for chapters (see below). To open the book at a different place, the user selects a new point along the groove and taps twice. The volume is opened at this new page and the working area is updated appropriately. The current maximum-sized volume of 900 pages represents the maximum resolution of the system. Clearly, when "opening" a document at a new page, the page that appears may not be exactly the one required, as when dipping into a real book. Thus a few pages may need to be turned after opening the document to get to the exact page. When a page is turned, the page number may be read out of the working area, if the pages are numbered.

For unpaged material, the same control works in a somewhat similar manner, but as a straight proportion of the total material, not quantised by pages, which could make finding the right place a lot less convenient . The second groove from the top provides a holophrast of the chapter headings and bookmarks in the book. This works very much like the holophrast columns for the working area, but the entities are chapters and bookmarks. Bookmarks will identify themselves, either using the default names generated when they are first inserted, or using names provided then or later by the user. Chapters may be expected to have names or numbers that can be used, and can also be read from the text being viewed in the normal way.

The third groove from the top provides a means of adjusting the window horizontally with respect to the material being viewed. If the lines of text are too wide to fit across the working area, they are initially pseudodisplayed with the leftmost character at the left edge of the window. Repositioning works very much like the vertical repositioning column described earlier. The window may be slid left and right across the material by running the finger along the third groove in the direction and by the amount that the user wishes the window to move, and then tapping once to signal readiness. Thus the movement ratio is one-to-one, and the change only occurs when the user is satisfied with the selection. When the window reaches the edge of the area defined for the material, it will beep and not move off the material. Normal pages will not require such movement, but it is provided to give flexibility in handling oversize page widths.

D. Host Access

The workstation may be used as a standalone device, providing many of the facilities expected of a personal computer. What is complete, so far, allows standalone editing. Interfaces to other people's packages, such as spreadsheets, have yet to be developed. However, it is also designed to provide enhanced terminal access to arbitrary host facilities. A standard 300/1200-baud modem, capable of answering or originating telephone calls automatically, allows data connections to be established to suitably equipped computers at remote locations. This includes public access information services. At this stage of development, the conveniences and advantages of using the Touch 'n Talk are rather dependent on the nature of the host software. However, a terminal emulator communicates with the host according to any of several standard terminal protocols, and updates the pseudo-display appropriately. Thus any software that can be operated using a normal terminal can, in principle, be operated by visually impaired people using Touch 'n Talk. Problems of optimized screen update, which uses control characters and text fragments to minimise the number of characters transmitted in updating the screen, are avoided, since the user only accesses the information once it is properly formatted on the pseudodisplay. Also, interruptions from time updates and the like do not occur. The user can determine the time by reading the appropriate portion of the pseudodisplay pad when it is convenient. Equally, multiplewindow operations are possible. The terminal emulator can also equate a page larger than the size of the pseudodisplay with the screen of the emulated terminal, and thus emulate a screen of arbitrary size. The window-dependent mechanisms discussed above will provide adequate means of viewing all parts. What we have not solved are the problems of drawing the user's attention to incoming error messages, unexpected screen updates, and worst of all, small changes. With a multiple window view for the sighted user, visual feedback, even in peripheral vision, tends to make even the smallest changes noticeable, without necessarily being distracting. For the visually impaired user, either every change must produce an auditory signal, which could be annoying, and still leaves the user with the problem of determining the full extent of the change(s), or the changes can be ignored, which could be worse. Error messages are an especial problem. Since the terminal emulator will pass "bell" characters as appropriate noises, host software can tag important updates by sounding the bell. However, the user still has to find out the full extent of the changes, which could range from a single character to the entire screen. One reasonable compromise is to beep once for every continous string of characters received, update the screen, and keep a secondary pseudodisplay showing only the most recent changes, with a facility for reading them both in and out of the whole screen context.

This approach would provide an automatic "difference" indicator between an older screen and the current screen; a secondary pseudodisplay geared to change information. The user could control the rate at which the older screen was updated (which would involve the Touch 'n Talk terminal emulator keeping auxiliary history data), and thus control the amount and currency of the change information displayed on the secondary display. The idea could obviously be elaborated to make finding changes of any reasonable age fairly straightforward.

Soft functions can be provided to enable and disable the change beep (which would not sound the same as the bell character), to switch between primary and secondary pseudodisplays, and to control the context in which changes are placed. When reading changes "in context," speech resources of intonation, rhythm, voice quality, and the like could be used to distinguish the context from the changed material. Coupled with suitable bell characters from the host, the user is in a strong position to judge the likely importance of changes (given knowledge of different current activities), and to find out exactly what they are, quickly and conveniently. This is the line that is being taken in the current version of Touch 'n Talk. A better solution will depend on either considerable applications specific intelligence in the workstation, or specially written software in the host that makes use of extended functions in the workstation. Other secondary pseudodisplays could be used to provide ready access to other types of information (such as local system information), within reason.

111. INITIAL EXPERIMENTAL EVALUATION

A. Introduction

As part of the initial work on this project, to gain confidence that the approach proposed was viable, two simple, related experiments were carried out using a working prototype of the device described above. The object of the experiments was to find out if the basic ideas would work in practice, and if so, whether there was any performance advantage for the new system compared to existing systems. A comparative experimental form was chosen to get around some of the difficulties inherent in absolute performance measurement. The tasks were chosen to resemble those encountered in real working situations, with menus, text access, and the like. However, rather than attempting to compare measures of overall performance in uncontrolled real work sessions, with a user satisfying arbitrary goals, which would raise many problems and involve many factors besides those of interest, selected controlled subtasks were chosen. The subtasks were designed to focus on the use of: pointing; structure; location determination; and direct manipulation. These were the facilities we were interested in transposing to a nonvisual domain. We were also interested in a comparison of sighted users wearing blindfolds and blind users. It was considered quite likely that the latter would not have the same spatial abilities as the former, which could adversely affect their performance relative to the blindfolded, normally-sighted subjects in trying to use spatial cues. However, blind subjects are in relatively short supply, and should not be imposed upon. They should therefore be reserved for necessary crucial tests, if normally sighted subjects can provide the same valid data in routine experiments. Thus the initial evaluation was crucial in this latter sense, as well as in the feasibility sense.

B. Apparatus

Two systems were used. One was a conventional key-operated talking terminal that resembled, in simplified form, typical current devices (the key device also referred to as the keyboard or even just KB). The other was a simplified version of the Touch 'n Talk, as described above (the pad device, also referred to as the speechpad or even just SP). Some effort was made not just to avoid biasing the experiment in favour of the new device but, if anything, to bias the experiment in favour of the old-style key-controlled talking terminal. We felt that if we tried very hard to tailor the key device to task characteristics appropriate to the intended use, as well as making it comparable to the pad device, and we still found a performance advantage for the pad device, we were more likely to be dealing with a real advantage rather than an artifact. Thus, although conventional talking keys on one current talking terminal resemble Fig. 3(a), the key layout actually used for the key-controlled device in the experiment is shown in Fig. 3(b). For comparison, a rather different, more complicated approach on another current key-controlled talking terminal appears as Fig. 3(c)

For the experimental layout, a help key accessed by the nonpreferred hand was also provided, in case subjects had difficulty remembering the effect of the keys. Adding the "up" and "down" keys allowed columnar data to be handled more easily ("previous line" and "next line" also reset the user cursor to start of line, whereas "up" and "down" did not). Columns are easily scanned with the pad

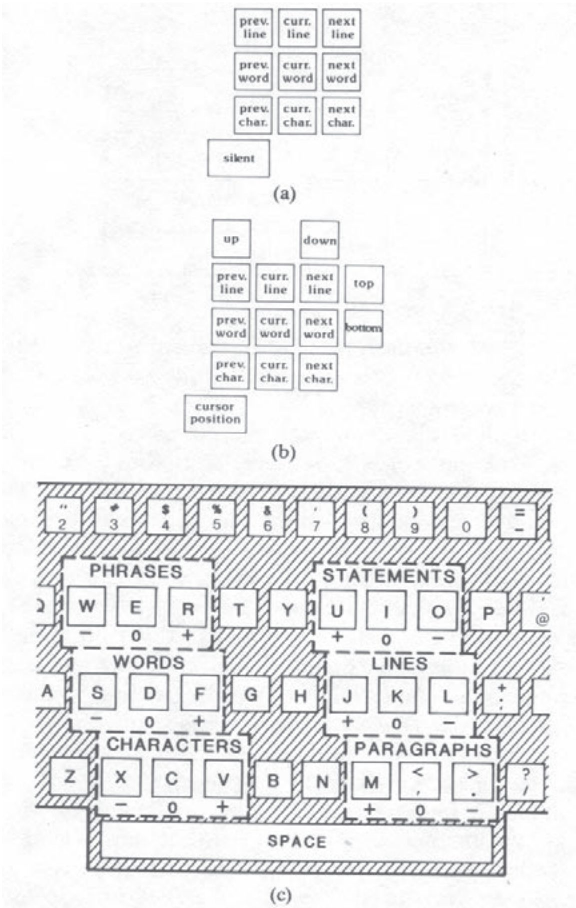

Fig. 3. Three forms of keypad to control speech on talking terminal. (a) Maryland Computer Systems "Information Through Speech."2 (b) Enhanced speech control pad used in the experiment. (c) Alternative way of controlling speech functions in talking terminal. (Note: Chording is required to resolve dual use of keys [1].)

device, so the addition of these keys to the key device was intended to reduce the advantage of the pad device in this respect. Adding the "top" and "bottom" keys reduced the advantage enjoyed by the pad device in getting to the top or bottom of the screen (important, since these are favored locations for display of frequently used and important information). Adding the cursor position readout key seemed an obvious refinement and again, in the context of the experiment, tended to reduce the purely functional advantage of the pad design, in contrast to the supposed perceptual/proprioceptive advantages that were at issue. Finally, instead of a key to silence output for the key device, we arranged that speech output only occurred as long as the key was held down. This was more natural, and analogous to the pad device. If the key was only depressed momentarily, so that no speech output was generated, the user cursor moved to the next unit, with only a beep to provide confirmatory feedback. Auto-repeat was provided to access successive units, except for speaking lines.

An existing commercial text-to-speech translator was used for both devices under test (the one built into the Maryland Computer Systems Information Through Speech[2] (ITS) system). This does not represent the state of the art, but was

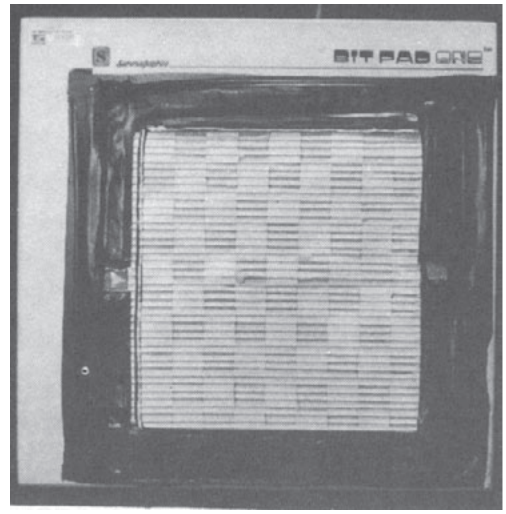

Fig. 4. Textured pad used in experiment

adequate. Its most serious limitation for our purposes was its lack of dynamic speech rate control. This acted to the disadvantage of the pad device for obvious reasons-speech rate and finger movement were somewhat disconnected. For the current implementation we are using our own system for text-to-speech conversion.

The pad device only had the line holopohrast column and the cursor finding column, apart from the working area functions. We felt a "help" key was not necessary—a feeling that was justified by the results. The textured pad was a rough prototype of the one in current use. A photograph of this prototype appears as Fig. 4 .

Both devices (pad and key) were implemented with most components in common, based on a Corvus Computer Systems Concept[3] Rather than a pure touch pad, a Summagraphics Bitpad One[4] served to sense finger position in the pad configuration, because there were (and still are) difficulties obtaining a suitably large, high resolution touch pad capable of operating with a textured surface or overlay. A suitable search coil was attached to the finger to simulate the stylus. This seemed to be a defect of our equipment at the time, but provided an unexpected insight in the end, an outcome at which we have already hinted. Data collection, including timing, was automatic, the raw results file being built up on an attached host. Raising and lowering the finger was detected because the range of our finger mounted coil was limited so that the digitizing signal vanished when the finger was raised.

C. Method

Each subject was brought in for three approximately one hour sessions. The first hour was devoted to pre-trial procedures: consent forms were signed; a pretest question-

[2] InformationThrough Speech is a registered trademark of Maryland Computer Systems.

[3] Concept is a registered trademark of Corvus Computer Systems Inc.

[4] Bitpad One is a registered trademark of Summagraphies Inc.

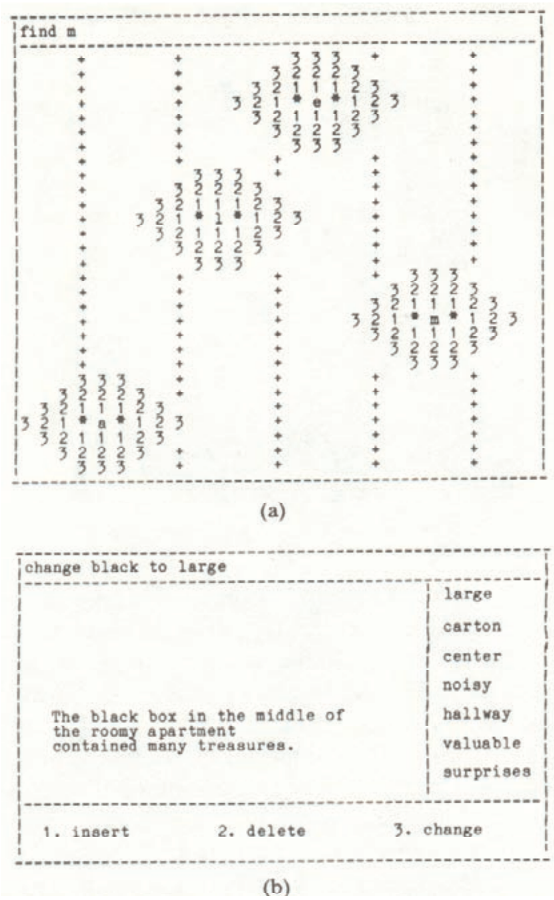

Fig. 5. Material layout in experimental tasks. (a) Illustration of material layout for target tests. (b) Illustration of material layout for edit tasks. (Note: Original displays had 39 rows in experiment)

naire was filled in; and subjects took a typing test and a T-maze running test to provide a measure of control for typing and spatial ability. The results were used in allocating subjects to cells, and reviewed later, to avoid the possibility that we might, by chance, put all those who were "fast" or had superior spatial ability, in one treatment condition. The remaining two sessions were devoted to the experimental tasks.

Of the two task types that formed the subject of the experiment, one was a game, based on targets and an ability to gain a feel for the spatial layout of the targets (used in session 2); and the other was a simplified screen editor that used a command menu (used in session 3). Figs 5(a) and 5(b) show, diagrammatically, typical screens for each task type.

Experiments to test relative speed of access to targets for different pointing methods have been carried out for the sighted (e.g. [3]). A target pops up in a random location, the subject keeps the target in visual focus, and uses some mechanism to move a cursor to the location. In this experiment, there was obviously no way of communicating such a system defined location directly to the user. Thus the locations had to be defined beforehand, and remembered (which is also consistent with real tasks), with the system asking for reference to any of these locations at random. Since we were not interested in the confounding effect of short-term memory limitations, the total number of locations accessed was kept within the general

TABLE I
EXPERIMENTAL DESIGN

Keyboard first	Material A first	S1, S3, S8, S10, S13
	Material B first	S6, S12, S15, S17
Speechpad first	Material A first	S5, S9, S11, S14
	Material B first	S2, S4, S7, S16

limits of short-term memory [8].

For both sessions 2 and 3 the overall procedure was the same. First the subject gained familiarity with the particular display. We refer to this as the exploration phase. The exploration phase lasted until the users felt they had mastered the device, in the sense that they did not have to think about how to use it, but could concentrate on the task. This took ten minutes or so and did not seem to be a major issue. Then the user was asked to find single items (we refer to this as the pretest). Finally, the user was asked to find combinations of items, in order (we refer to this as the main test). The pretest was intended to mediate between the user-determined exploration phase, and the main test, to make sure that the user "knew" the different locations in some real sense before performing the main task. Thus there were really four test series for each device on which measurements were taken. A target pretest, a target main test, an edit pretest and an edit main test. As there were two devices, each subject completed a total of eight test series. As a precaution against the effects of the material used confounding the result, some subjects used one set of test material (material A) and others used another (material B). As a precaution against order effects, half the subjects used the key device first and half the pad device. Thus the experimental design was a mixed factors design with two grouping variables or between subjects factors (first device: keyboard versus speechpad; and first material: material A versus material B), and one within subjects (or repeated measure) factor (device: keyboard versus speechpad), as illustrated in Table I.

Each test series consisted of six, seven or eight trials (depending on the test series concerned) using the screen with which the user had just become familiar. Each trial started by having the user read a machine generated description of the task from the top line of the screen. The task description was continuously available, and the number of times a user found it necessary to re-read the task description was assumed to be a useful measure of the mental load in using the device (pad or key) to complete the task. Appendix A provides copies of the actual instructions to subjects, together with the pre- and post-test questionnaires. The instructions were actually recorded on cassette and played to subjects, for obvious reasons, and the experimenter wrote in questionnaire answers for the blind subjects.

The same pattern was repeated for each device, for each task, the starting device being assigned at random within the constraints imposed by the experimental design. Thus a given subject would execute four tasks on each device, each task comprising a number of trials, would start with either the key device or the pad device, and would use one of two types of material.

Six dependent variables were measured. In each trial, the subject signified satisfaction with each location selection by saying "mark," the experimenter acting in the role of speech recogniser. Thus time could be measured for each trial from starting to read the task description to completing the task, whether single item selection or a sequential task. Incorrect selection led to the generation of an error message and signified an error. Errors were also counted. In addition, the number of words listened to (words), the number of reversals of scanning direction (implying the subject overshot an intended target, was learning structure, or was somehow lost) (reversals), the number of readings of the instruction line (task), and number of times the help key was used (for the key device) were all counted. In practice, the help key was simply not used, and there were no data to analyze. Thus five dependent variables appear in our analysis .

After the last task had been completed, subjects were asked to fill in a posttest questionnaire. Additional informal data were collected by the experimenter, who made notes on the progress of the trials and any remarks made by subjects. These form part of the basis for some of our interpretations.

D. Subjects

Twelve sighted students (subjects 1 to 12) and five blind students (subjects 13 to 17) acted as subjects. All were volunteers and were paid $4 an hour. All but one of the subjects were students at the University of Calgary, ranging in age from 20 to 35. All the sighted students were majoring in computer science and therefore had at least two years of core courses in computer science. A large proportion of them were drawn from the Computer Science 481 (Human-Computer Interaction) course. None of the blind students was majoring in computer science, but two of them had taken one computer science course each.

E. Null Hypotheses

The following null hypotheses were considered. In each case, the comparison is between performance with the pad device and performance with the key device.

- I There is no difference in time taken to complete tasks
- II There is no difference in the number of errors made
- III There is no difference in the number of words listened to
- IV There is no difference in the number of times the task instructions are read
- V There is no difference in the number of scanning direction reversals
- VI The two devices are equally easy to learn and remember

The rationale for some of the above is also of interest. Time and errors are traditional performance measures (I and II). For III, we assume that, in order to perform the task, the user only needs to read a certain number of words. Any words beyond that limit should be unnecessary, and only generated due to some kind of device handling problems. If one device consistently requires a subject to listen to more words than the other, it could reasonably be considered inferior. For IV, we assume that if the subjects have to re-read the definition of a simple task, there has been some interference due to the cognitive load in using the device. If one device consistently causes users to read the task description more often than the other, it could reasonably be considered harder to use and therefore inferior. In V, we may be concerned either with a degree of flexibility in searching (more reversals suggests greater flexibility) or with a degree of disorientation (more reversals suggests more confusion). The hypothesis was exploratory. Finally, for VI, it ws hypothesized that if the key device were as easy to learn and use as the pad device, the subjects would not need to use the help key for that device, since its operation would be equally easy to learn in the initial stages, prior to testing. One needs little help in operating a finger, if the pointing model has any validity at all.

F. Results and Analysis

The initial statistical analysis was carried out using the SPSS statistics package (SPSS for Multics/6880, Version H. Release 9.1, 1982) (see also [9]). However, a complete re-analysis of all the results was carried out as a check, using the BMDP package [4], especially the analysis of variance with repeated measures subprogram *bmdp2v*. The BMDP package proved much nicer to work with and is the source for nearly all the figures quoted (with the exception of the initial test for differences between blind and sighted subjects). Where the analyses were comparable, however, the two sets of results were in close agreement. The main analyses used the subject means for the dependent variables from the four test series as input data. These results appear in Appendix B, together with the responses to the pretest and posttest questionnaires, except that since help was essentially not used, no data for help are presented.

The typing and T-maze tests produced widely varying results that were well distributed over the different cells in the experimental design. No recourse to explanations based on anomalies in these results was needed, and they are not considered further.

The main performance measure for this experiment was time taken to complete tasks. Before proceeding further, a preliminary analysis, breaking the trials down by task type, and sightedness of subjects, was done to see if there was a compelling reason for distinguishing the blind subjects

TABLE II
RESULTS OF PRELIMINARY t-TEST ON THE DIFFERENCES BETWEEN SIGHTED AND BLIND SUBJECTS[a]

		Mean time per trial								Test of variances		Sighted versus blind trial means		
		Sighted subjects				Blind subjects								
Test	Device	Mean Time	Variance	N	kb/sp ratio	Mean Time	Variance	N	kb/sp ratio	F	p(F)	t	dF	p(t)
Target pretest	kb	1290.0	448190.1	12	1.82	1370.2	411161.8	5	1.84	1.09	1.000	0.23	15	0.823
	sp	708.6	51331.2	12		746.6	79622.2	5		1.55	0.510	0.29	15	0.773
Target main test	kb	2105.8	404025.5	12	1.25	2451.4	456669.1	5	1.64	1.13	0.783	1.00	15	0.331
	sp	1692.1	277400.2	12		1495.2	569080.1	5		2.05	0.312	0.62	15	0.544
Edit pretest	kb	1100.3	156197.3	12	1.19	1004.6	274808.7	5	1.62	1.76	0.414	0.42	15	0.684
	sp	921.1	412420.8	12		620.3	71422.6	5		5.77	0.105	1.00	15	0.335
Edit main test	kb	2007.6	234512.6	12	1.20	2014.7	63032.1	5	1.69	3.72	0.216	0.03	15	0.976
	sp	1678.5	300734.9	12		1189.8	146745.7	5		2.05	0.510	1.80	15	0.092

[a]Time is measured in 50-ms units. Thus 1000 units = 50 s.

TABLE III
SUMMARY OF MIXED FACTORS ANALYSIS OF VARIANCE FOR DEPENDENT VARIABLE "MEAN TRIAL TIME PER SUBJECT"[a]

Source of variation	Sum of squares	Degrees of freedom	Mean square	F ratio	Significance of F ratio	Mean kb time / Mean sp time / Kb/sp time ratio
Target pretest:						
Device type	2853224.4	1	2853224.4	18.76	0.0008***	1316.6
Dev-type x 1st.-dev	249730.8	1	249730.8	1.64	0.2225	719.8
Dev-type x mat-ord	34842.1	1	34842.1	0.23	0.6402	1.83
Dev-type x 1st.-dev x mat-ord	71598.2	1	71598.2	0.47	0.5047	
Residual	1977369.5	13	152105.3			
Target main test:						
Device type	2701534.2	1	2701534.2	9.11	0.0099**	2207.4
Dev-type x 1st.-dev	63636.0	1	63636.0	0.21	0.6508	1634.2
Dev-type x mat-ord	123340.4	1	123340.4	0.42	0.5301	1.36
Dev-type x 1st.-dev x mat-ord	56806.7	1	56806.7	0.42	0.6687	
Residual	3853567.9	13	3853567.9			
Edit pretest:						
Device type	430137.1	1	430137.1	1.89	0.1928	1072.2
Dev-type x 1st.-dev	87337.4	1	87337.4	0.38	0.5467	832.6
Dev-type x mat-ord	55983.3	1	55983.3	0.25	0.6285	1.29
Dev-type x 1st.-dev x mat-ord	371095.8	1	371095.8	1.63	0.2244	
Residual	2964104.8	13	228008.1			
Target main test:						
Device type	1904482.3	1	1904482.3	11.45	0.0049**	2009.7
Dev-type x 1st.-dev	23564.7	1	23564.7	0.14	0.7127	1534.7
Dev-type x mat-ord	5273.4	1	5273.4	0.03	0.8614	1.31
Dev-type x 1st.-dev x mat-ord	64991.7	1	64991.7	0.39	0.5427	
Residual	2161823.4	13	166294.1			

[a]There are two grouping factors (device order and material order), and one within factor (device type). The table is based on 17 subjects and shows the four different tests. Time is measured in 50-ms units. Thus 1000 units = 50 s

**Significant at the 0.01 level.

***Significant at the 0.001 level.

from the normally sighted subjects based on their performance. Two factors were important in this decision: homogeneity of variance, and difference in group means. Table II summarizes the results of this analysis, which is broken down by test type. None of test conditions shows a significant inhomogeneity of variance (although there is a suggestive trend for the speechpad in the edit pretest). Likewise, a t-test of the differences between the test means for blind subjects versus sighted subjects failed to reveal anything close to a significant difference. It is interesting that there is a tendency for the blind subjects to do relatively better on the speechpad than on the key device, compared to the sighted subjects, as revealed by a comparison of the times, and the keyboard to speechpad time ratio. This suggests a greater ability to use the pad device on the part of our blind subjects. However, it is not statistically significant, and does not warrant separate analysis. Thus the results for all subjects were pooled.

The analysis of variance for the pooled results is shown as Table III. There is no evidence of interaction between the main effect and device order or material used, for any task. Three of the trial conditions (that is, all except for the edit pretest) show a significant (0.01 level) or highly significant (0.001 level) difference between the two devices. The magnitude of the difference is also quite significant in a non-statistical sense, since, on average, subjects using the key device took nearly half as long again to complete the tasks as those using the pad device. The average keyboard to speechpad time ratio over all trials was 1.45. One subject in one task completed the task 6 times faster using

TABLE IV
SUMMARY OF MIXED FACTORS ANALYSIS OF VARIANCE FOR DEPENDENT VARIABLE "MEAN TRIAL TIME PER SUBJECT," FOR THE EDIT PRETEST, BROKEN DOWN INTO TWO PARTS[a]

Source of variation	Sum of squares	Degrees of freedom	Mean square	F ratio	Significance of F ratio	Mean kb time / Mean sp time / Kb/sp time ratio
Edit pretest, trials 1 to 3:						
Device type	573705.3	1	573705.3	6.47	0.0245*	848.1
Dev-type x 1st.-dev	4274.2	1	4274.2	0.05	0.8297	581.5
Dev-type x mat-ord	1001.3	1	1001.3	0.01	0.9170	1.46
Dev-type x 1st.-dev x mat-ord	138446.7	1	138446.7	1.56	0.2336	
Residual	1153190.1	13	88706.9			
Edit pretest, trials 4 to 6:						
Device type	212669.0	1	212669.0	0.32	0.5814	1296.3
Dev-type x 1st.-dev	130266.4	1	130266.4	0.20	0.6654	1113.9
Dev-type x mat-ord	21334.1	1	21334.1	0.03	0.8606	1.16
Dev-type x 1st.-dev x mat-ord	882417.6	1	882417.6	1.33	0.2702	
Residual	8648003.3	13	665231.0			

[a]The two parts being trial 1–3 and trials 4–6. There are two grouping factors (device order and material order), and one within factor (device type). The table is based on 17 subjects. Time is measured in 50-ms units. Thus 1000 units = 50 s.
*Significant at the 0.05 level.

TABLE V
MAIN EFFECTS SUMMARY ABSTRACTED FROM ANALYSES OF VARIANCE FOR MEASURES OTHER THAN TIME[a]

Test	Measure	Keyboard		Speechpad		Degrees of freedom	F ratio	Significance of F
		Mean	Variance	Mean	Variance			
Target pretest	Errors	0.1	0.1	0.1	0.0	1	0.25	0.6224
	Words	32.6	1096.8	29.9	558.9	1	0.30	0.5958
	Task	1.2	0.1	1.3	0.1	1	4.25	0.0598
	Reversals	2.7	16.9	7.1	48.9	1	23.01	0.0003***
Target maintest	Errors	0.2	0.0	0.2	0.1	1	0.74	0.4058
	Words	59.1	756.4	59.2	1219.9	1	0.01	0.9102
	Task	1.7	0.3	1.9	0.9	1	0.66	0.4304
	Reversals	7.1	23.5	18.3	122.0	1	31.49	0.0001***
Edit pretest	Errors	0.2	0.0	0.2	0.0	1	0.02	0.9013
	Words	32.9	576.0	22.2	453.9	1	8.65	0.0115*
	Task	1.2	0.0	1.2	0.0	1	0.00	0.9501
	Reversals	2.0	6.4	8.9	195.7	1	19.10	0.0008***
Edit maintest	Errors	0.3	0.0	0.3	0.3	1	0.19	0.6690
	Words	60.8	946.4	40.9	1254.1	1	22.24	0.0004***
	Task	1.6	0.3	1.5	0.2	1	2.42	0.1440
	Reversals	6.7	14.4	13.9	115.5	1	34.07	0.0001***

[a]The residual had 13° of freedom in each case.
**Significant at the 0.01 level.
***Significant at the 0.001 level.

the pad device, despite a respectable time using the key device (blind subject number 13, in the edit pretest, had a mean time using the keyboard of 1918.5 time units (95.9 s), and a mean time using the speechpad of 311.5 time units (15.6 s), giving a ratio of 6.16).

The edit pretest does not show a significant difference between the mean completion times, even though the keyboard to speechpad ratio is comparable to (though slightly lower than) the other test condition ratios. Splitting the edit pretest into two parts, trials 1-3 and trials 4-6 shows that the difference for trials 1 -3 is just statistically significant (at the 0.05 level) (see Table IV). The only comment that can be made, at this point, is that the residual error in the data was proportionately greater in the edit pretest than in the other tests (hence the lower value of F despite the difference in means) and that we thought this may have been due to the inhomogenous nature of the edit pretest. That is, in trials 1-3, subjects were asked to find a word in the sentence, which could be readily learned on first hearing it, whereas in trials 4-6. they were asked to find a word from one of the other two areas, which were probably less easily dealt with, having fewer constraints. Thus potentially differing stategies may have affected the data differently, leading to the greater residual, and allowing statistical significance to be detected on the split analysis. The actual results for the split analysis do not provide very strong support for this suggestion. Perhaps the more complicated display simply led to a wider variety of strategies. It is worth noting that a split t-test of the difference between sighted subjects and blind subjects, based on trials 1-3 versus trials 4-6 in the edit pretest, still showed no reason to analyze them separately. These data are not tabulated.

The other dependent variables were also subjected to a repeated measures analysis of variance. The results for the main effects are summarised in Table V. In all cases there were significantly more reversals in scanning direction (0.001 level) when using the pad device than when using the key device. For the edit task, significantly more words were read when using the key device than when using the pad device (0.05 level for the pretest and 0.001 level for the main test). No other differences based on main effects were significant. Three interactions were statistically significant (device with first device for errors on the target pretest, and device with first device and material order for task line readings on both the edit pretest and the edit main test). These interactions are ignored partly because of small absolute significance (especially for the occurrence of errors—there being a slightly greater chance of error using the pad device if it were used first and a slightly smaller chance of an error using the key device if it were used first); and partly because they allowed no rationization (subjects were somewhat more likely to revisit the task line when using the pad device, if they used it first with material A on the edit pretest, but the opposite effect was observed on the edit main test).

G. Discussion

H I — There is no difference in time taken to complete tasks: Time was the main performance measure for our experiment. For all test series except the edit pre-test there was a substantial (45 percent faster on average) statistically significant (at the 0.01 level or better) difference in favour of the pad device regardless of the order in which the devices were used. For the edit pretest, a performance advantage was present, and this was statistically significant for trials 1-3 (at the 0.05 level). Varying the material had no effect on this finding. Lack of familiarity with key-based devices does not seem to be an issue. Our blind subjects were all competent, some skilled keyboard users, and all the sighted subjects were regular keyboard users. The keys were simple to use as evidenced by the almost total lack of recourse to the help key by subjects, and the relatively similar error rates, which were low for both devices. Thus we firmly reject the first (main) null hypothesis and conclude that not only is the new pad device usable for simple but realistic tasks, but also it may offer significant performance advantages.

This finding, that pointing is superior to keys for location determination and access, is consistent with studies using sighted subjects carried out on other pointing devices, comparing them to keys (e.g. [3]). What is significant in our study is that apparently similar advantages can be made available to blind users, based on an effective translation of the direct manipulation domain normally associated with tight visual feedback into the tactile/ proprioceptive/audio domain. It must be emphasized that this finding is preliminary, and based on fairly simple task elements. However, within the limitations of our experimental set-up, it seems fairly robust.

H II—There is no difference in the number of errors made: As already noted, we do not reject this hypothesis. There was no evidence of any difference and error rates were very low on both devices. Such errors as did occur were usually caused by subjects hearing the target word last, but moving onto adjacent white space before marking. This was regarded as an error, as it would cause problems in a real editing situation. Subjects soon learned to double check location before marking but clearly problems with synthesizer speech rate control, and a need to improve the nonspeech audio feedback are relevant here. Progress can certainly be made in these respects. It is worth noting that the error data suggest that the problem of time—error trade-off was not a problem in our experiment.

H III — There is no difference in the number of words listened to: This hypothesis is not rejected for the target task, but is rejected for the edit task. Many more synthesized words (48 percent more) were called out by subjects using the key device than by subjects using the pad device, and this difference was statistically significant (at the 0.05 level for the edit pretest, and the 0.001 level for the edit main test). That this effect was task dependent very likely reflects the highly artificial structure of the display for the target task, in which subjects were in effect compelled to listen to a good many words to determine their location relative to the target cell, regardless of the device used. In the editing task, the natural structure of the sentences, coupled with the richer structure of the display, could provide much more location information for each of the words read, and allow the subject to complete the task on the basis of fewer of them if they really gained a "feel" for the structure. It suggests that the pad device is superior in such a situation and is clearly an area for more research .

H IV — There is no difference in the number of times the instructions are read: Although this hypothesis is not rejected for any task, some comment is in order. Observations made by the experimenter suggested that subjects using the keyboard were reluctant to return to the command line to check the task, preferring to take a chance on marking the wrong thing rather than lose the location just determined. Pad users were able to keep track of the location with the nonpreferred hand while referring back to the command line with the hand holding/ wearing the stylus. Precisely what effect this behavior had was not directly measurable within the experimental design and, in any case, as already noted, there were few errors with either device, so the keyboard users' strategy was obviously a good one. Nevertheless, a useful addition to the key device would seem to be a user-cursor memory, requiring a save and restore key, which would overcome the problem. However, this illustrates how quickly a simple key approach can become too complicated, if useful facilities are added indiscriminately. However, with the flexibility and character of the pad device, some of these facilities not only come "for free" in the sense that they do not require special hardware provision, but are also very natural, which could alleviate the potential problem of excess cognitive complexity arising from providing too many facilities with inadequate integration. It seems obvious that, with extended use, a system that does not provide easy confirmation will engender greater anxiety and fatigue, and probably cause a degradation of performance. Thus, although it is not justified directly, our feeling is that the pad device was superior in providing easy cursor memory, and that an equivalent function for the key device would not completely redress the balance. However, some of this advantage might evaporate if users were also using a keyboard to enter text or other data, forcing tactile discontinuity [6]. This is why it is important to be able to use more than one cursor, and manipulate them easily and naturally, whatever the device. This kind of consideration typifies the distinction between providing required functionality, and providing that functionality in convenient, usable form. Care in thinking through and testing details of an interactive system at this level (and even lower) is crucial to making excellent computer-human interfaces. It is also worth noting that a true touch device might preclude the use of the nonpreferred hand for marking unless the device were capable of resolving multiple contact points. Even then, suitable algorithms would need to be developed to handle multiple contacts intelligently. This could prove an interesting research problem, given that elbows, coffee cups and the like might all have to be distinguished as well. Given our experience with the use of a stylus, rather than touch, however, it may not be of practical concern (see below).

H V — There is no difference in the number of scanning direction reversals: Typically, subjects reversed scanning direction from two to four times more often when using the pad device than when using the key device. The differences are all significant at the 0.001 level or better so that we firmly reject the null hypothesis. Since performance was clearly superior using the pad device, with low error rates on both devices, we believe that the difference again reflects the flexibility of the device, rather than some kind of problems faced by subjects. To reverse direction on the keyboard required subjects to change keys while reversing on the pad presented no such discontinuity. We speculate that the ease of reversal may well be part of the reason for the superiority of the pad device in our experiment, allowing more effective searching with less cognitive load, although we have no direct evidence for this Certainly, such a factor would promote structure determination. As it happens, we obtained some direct evidence which we have interpreted in support of the idea that the pad device promoted structure learning while the key device did not. The argument is presented in the section following discussion of the null hypotheses .

H VI — The two devices are equally easy to learn and remember: The help key for the key device was used so rarely that the data generated were insignificant in an absolute sense, and are not included in the tabulations. This suggests that the key device could be completely mastered during the first exploration phase. This was also true of the pad device, which had no help key. Thus we have no reason to reject the null hypothesis. However, it must be remembered that the key device was a greatly simplified version of a full talking terminal, which may quickly become complex through the addition of needed ancillary functions that may not require special provision in a pad device. Thus a key device set up for more complex situations may require a well supported

TABLE VI
AVERAGED RESULTS FROM THE POST-TEST QUESTIONNAIRE[a]

Detailed question	sighted	blind
1. The keyboard is hard to learn	- 1.04	- 0.6
2. The speechpad is hard to learn	- 0.95	- 1.4
3. It is easy to become confused using the speech-pad	- 0.4	- 1.2
4. It is easy to become confused using the keyboard	- 0.375	0.0
5. I was usually able to complete the task more quickly using the keyboard	0.0	- 1.2
6. I was usually able to find the place I wanted more quickly using the speech-pad	0.2	1.4
7. Which device was easiest to learn	0.29 (s)	0.6 (s)
8. Which device caused most confusion	-0.09 (k)	-0.8 (k)
9. Which device did you prefer using	-0.29 (k)	0.4 (s)
10. Which device would you recommend to a friend	-0.09 (k)	0.6 (s)
11. If you had your choice of only one device, which one would you choose?	-0.09 (k)	0.6 (s)

[a] The presentation of questions 1–6 was in the form of a linear scale, as shown in Appendix A.

help key. Our intuition is that it would. Even the pad device may require help facilities in a full implementation. We are certainly incorporating help facilities in the workstation based on our new design, although it is primarily intended to call up systems and applications oriented help. For additional information on learnabihty and ease of use, we turn to the answers to the post-test questionnaire in Table VI (note: the individual questionnaire results appear in Appendix B, Table B-V). Questions 1-6 of that questionnaire elicited a scaled response on a range from - 2 (strongly disagree) to + 2 (strongly agree), and the tabulated results show the average scale response. We tried to incorporate a degree of internal validation in the questionnaire by eliciting similar information in two rather differently phrased questions for each major point of interest.

General: Neither type of subject reported finding either of the devices hard to learn (Questions 1 and 2). On average, they disagreed with the suggestion that either was hard to learn (scale - 1), although the blind subjects tended to disagree less strongly with that suggestion about the key device (scale = -0.6). Questions 3 and 4 were concerned with ease of use, specifically confusion in use. The results are less clear cut, but there is certainly a difference between the two types of subject. Normally sighted subjects did not feel either device caused confusion, with roughly equal strength (scale = -0.4). The blind subjects (who were our real target user population) were ambivalent about the key device (scale 0.0), but disagreed that the pad device caused confusion (scale - 1.2). Finally Questions 5 and 6 probed the subjective impression of effectiveness for the two devices. Here the difference between the two types of subject is even more marked. The normally sighted subjects did not have strong feelings either way as to which device was faster in use for location determination, even though objective measurement showed that they were faster with the pad. The blind subjects, however, disagreed with the suggestion that they found the key device faster (scale = -1.2) and agreed with the suggestion that they found the pad device faster (scale = 1.4). These subjective impressions were borne out by the objective results.

The remaining questions were really repeated requests for similar information, rephrased and in different response form, to provide a built-in check on the subjective data, as already noted. On the whole they are consistent with the results already discussed. It is interesting to note that even though the sighted subjects somewhat preferred the key device (Questions 9, 10, and 11), they found the pad device easier to learn and found it caused no more confusion (Questions 7 and 8). Also, of course, their objective performance was very much better on the pad device and, in fact, not significantly different to that of the blind subjects. We can only assume that, as they were all frequent keyboard users, the apparently paradoxical preference for the key device somehow reflected their great familiarity with keyboards, compared to special devices with textured surfaces, which they had almost certainly never encountered before. Based on our own (unpublished) experience in other experiments involving blindfolds, we feel that this effect may have been enhanced by the stress of wearing a blindfold. Blind subjects, on the other hand, are used to feeling and exploring without seeing as part of their everyday activities, including (for some) reading Braille books. It seems that the pad device format has potential for blind computer users in the sense that not only does it work, but it is also immediately acceptable and comfortable. It also illustrates a cautionary note in further experiments. While we may assume that performance may be measured in similar tasks using sighted subjects wearing blindfolds, as a basis for design, the subjective aspect of users can only safely be investigated using blind subjects. Subjective preference does not necessarily follow conditions for optimum performance. In any case, frequent revalidation of normally sighted versus blind subjects must be undertaken, especially if the aspects being investigated, or the format of the experiment, were to differ significantly from our study.

It should also be noted that the context for blind subjects in answering the questionnaire was quite different from that of the normally sighted subjects. This, we speculate, is part of the explanation for the differences in response. All of the blind subjects had used some form of talking terminal before, and, for them, this experiment presented a real situation under real conditions. For the sighted it was more like some strange sort of game. Thus the blind subjects were far more concerned with the usability of the devices. Given their background, the normally sighted subjects, although very familiar with keys, had probably had little or no experience with pointing devices, and certainly had none using such a device without the benefit of sight, as already noted. Although we specifically asked subjects about their experience with talking terminals (in the pretest questionnaire), we overlooked the almost equally relevant question about experience with pointing devices.

There are a few other points that are worth discussing. First,

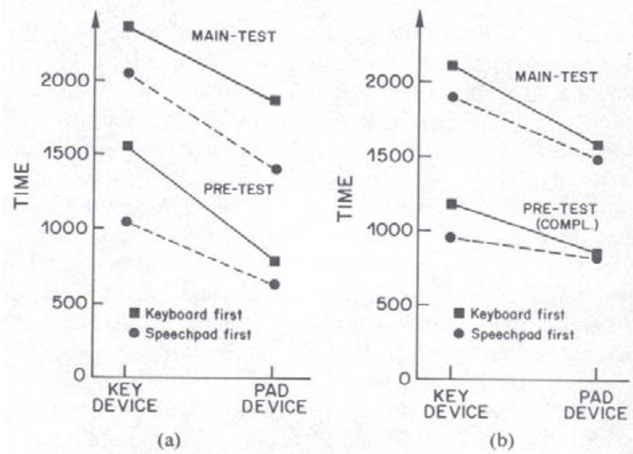

Fig. 6. Mean trial times for two devices broken down by test and by test and by groups starting with different devices. (a) Target test mean trial times. (b) Edit test mean trial times. (Time in 50 ms units.)

it was stated earlier that additional direct evidence concerning the value of the pad device as a structure visualising aid had come to light. Figs. 6(a) and 6(b) illustrate graphically the device order interaction with mean time to complete a trial under various conditions. As previously noted, the order interaction was not statistically significant, as is also suggested by the graphs. However, the graphs do reveal a consistent effect that should be investigated further.

There were two groups: subjects 1, 3, 6, 8, 10, 12, 13, 15, and 17 who used the key device first in all tests; and the remainder who used the pad device first. During trials with different devices, either learning occurred on the first device used, that could be transferred usefully to trials using the second device, or it did not. If such learning did not occur, then one would expect the mean trial times for a given test type and device to be roughly equal for the two groups. This is only obviously true for subjects using the pad device for the edit pretest. If on the other hand useful transfer did take place equally for either device, from the first device used to the second then one would expect the mean trial times for a given test type and device to be shorter for the group using the device second than for the group using the device first, for either device, since the group using either device second would have the benefit of useful learning transferred from their experience of using the other device first. In fact, as is clear from the absence of significant interactions and the direction of the main effect, this effect only occurred for the keyboard. That is, subjects using the key device first were slower on the key device than subjects using the pad device first, but they were also no faster on the pad device than subjects using the pad device first. They were usually slower (except for the edit pretest), the extreme case occurring for the target main test. We interpret this as suggesting that something could be learnt using the pad device for a test series that could he transferred to subsequent use of the key device, but that the reverse either did not happen, or happened to a much lesser degree. An elaboration of our argument follows.

The mean total elapsed time for completion of the target type trials using the key device was considerably longer than the mean total elapsed time using the pad device. The difference was less extreme for the edit type trials. Thus one effect taking place may be fatigue. Those subjects using the key device first took longer and would therefore have been relatively more fatigued when faced with the second device (pad), than those using the pad device first were when faced with the key device. Such a hypothesis is consistent with our informal observations during the experiment. However, fatigue cannot explain why subjects who started on the pad device performed better on the key device than subjects who started on the key device. We therefore hypothesize that considerable transferrable learning takes place when subjects use the pad device-for example, in terms of forming some usable mental representation of the structure of the test material, but that such learning does not take place when using the key device. This is consistent with our initial intuitions when developing our approach. Further research is clearly called for.

The second insight, which ran counter to our intuitions, concerned the use of a stylus. We had considered this to be an unfortunate compromise in setting up our experiment, and took pains to make the finger the main focus as a locator, for what seemed obvious reasons. In fact, the sensor attached to the finger seemed to cause both problems and irritation (as might be expected) but a number of subjects actually preferred to remove the sensor attachment from their finger and use it as a normal stylus while, at the same time, feeling the surface with the finger near the stylus. They preferred this arrangement. Discussions with subjects following the experiment revealed that they felt much more in control when using the sensor as a normal stylus because they could still feel the features of the surface, but knew precisely which point was indicated because of their awareness of the stylus tip position. It seems that, even with a "proper" touch surface, the finger would still provide only a rather imprecise feeling for the physical location. Clearly further research is needed but, for now, we are proceeding with a normal stylus, in conjunction with the Bitpad One and overlay, as shown in Fig. 1.

Finally, we turn to a minor anomaly of the individual data. The three fastest subjects on the pad device were all blind subjects (subjects 13, 14 and 16, as judged by overall mean trial time). The other two blind subjects turned in completely typical times (seven sighted subjects were faster, five slower). All that distinguished the three fast blind subjects from the other blind subjects to our knowledge was the relatively early age of onset of their blindness. Whether or not this led to better development of their tactile, proprioceptive, and spatial-memory abilities, or whether these subjects were simply exceptional as evidenced by their reaching university level despite their lifetime disability is yet another possibility for some interesting follow-up research.

IV. CONCLUSION

In this paper we have presented a design for an interactive direct-manipulation device intended to improve computer access for blind computer users, and make it possible for them to use modern systems that often depend critically on pointing and selection based on structured displays. It could easily be adapted to iconic displays, though without the iconic advantages enjoyed by normally sighted users (we wonder if auditory equivalents of visual icons could be developed and automatically associated with their visual counterparts to regain these advantages). We have also presented the results of an experiment intended to validate the basic principles of the device. Besides providing data about the device, the experiment embodies an approach to evaluation of interface components that has been used by others. Design and experiment are iterative. The current design, detailed in the first part of the paper, represents the second iteration, following the initial design and the conduct of the experiment.

We believe that our design represents a workable initial solution to the problem of substituting for a restricted visual channel in multimodal human-computer dialogue. A personal computer equipped with a digitizing tablet, as well as a speech synthesizer and other equipment, provides the basis for a special workstation designed to allow the visually impaired user access comparable to that now available to normally sighted colleagues. Touch, proprioceptive feedback, and voice feedback, coupled with the detection of natural gestures, like page turning movements and reversals, may be used to avoid the need for a multiplicity of special functions accessed by complicated keybased devices, and may be used to allow a spatial image of structured screens to be exploited without the need for visual feedback. Experimental evidence suggests that the basic device, in simple task situations: offers close to a 50 percent performance advantage over conventional talking terminals; is perceived to be easier to learn and use; and is preferred by the target population of blind users. It also suggests caution in establishing subjective preference, as opposed to objective performance measures, since the subjective preferences of the sighted users seem to differ from those of blind users, and run counter to what might be expected from objective performance measures.

A prototype workstation, based on the design presented, and the experiments performed, is near completion at the University of Calgary.

APPENDIX A[5]
INSTRUCTIONS FOR PARTICIPANTS IN THE EXPERIMENT

A. General Instructions

Thank you very much for volunteering to take part in this experiment. In the following, we will give you some general information about the experiment and its purpose,

[5]Note: Subjects were presented with this information in spoken form.

the procedures which are to be followed, and your rights as a subject in an experiment.

Please be aware at all times that you may discontinue your participation in this experiment at any time. You are under no obligation to complete the whole procedure. After you have read these instructions, we would like you to sign a declaration that you have been given complete instructions which you have understood, that you have volunteered for this experiment and that you understand that you may stop and leave at any point in the experiment.

Now some more information about the experiment itself. For a larger research project, we are testing different means of accessing a fixed display of text, when the only way of finding out what is displayed is to hear the words spoken. This is a situation with which visually impaired computer users are confronted. Thus, you will be blindfolded for the whole experiment. If you don't like to be blindfolded, please tell the experimenter right away. In that case, you probably should not take part in the experiment.

In the experiment at hand, you will get to use two different devices: In one set-up, you will be provided with a number of keys which move an imaginary pointer around on a page-like area on which words are arranged. As this pointer reaches a word, the word is spoken. With the other device, you will touch a surface which has embossed lines on it, and you will use your own finger to point at words . The purpose of the experiment is to determine if either of the devices enables you to find and use some of the words. We will be measuring the time taken from start to finish in doing this, as well as the number of errors made. However, it is the devices we are testing, not you. We expect you to have some difficulty. This is necessary to allow us to judge the devices and the experiment has been designed to provide a challenge for this reason. However, please work as fast and accurate as possible so that we can compare the devices properly.

During the experiment, for each tool, you will be introduced to it first, and will be given some time to play with it and get acquainted with it. You will then be presented with instructions describing the task you are to perform. There are two different tasks, a target-finding game, and an editing task. In both tasks, you will be doing similar things. In the beginning, you will be asked to familiarize yourself with the lay-out of the page on which the words are laid out. (Note that you cannot feel the words, but in the touch task you can feel the lines.). You can take as long as you like, within practical limits. When you are ready to proceed, you will be given a task, like finding a word, letter or number on the display. Please, find this item, and mark it, by saying "MARK." In this experiment, we do not pursue the "normal" way of marking things through the pressing of buttons, because this might disturb you in operating the device. This is why we are giving you a voice-recognition device. When you say "mark," the computer will understand this and try to mark the item at which you are pointing. You will then get an answer revealing whether or not the mark was set or whether you missed the item expected. You will be given some practice in order to get used to speaking to a machine.

So much for the general overview. In the first part of the experiment, you will only be taking a pre-test and answer some questions which we think are important for the interpretation of the results of the experiment. In the second part, you will be introduced to one of the tools, and you will perform the first task with it. Then you will do the same thing using the other device. In the third part of the experiment, you will then be presented with a slightly more complicated task. Again you will be asked to perform tasks on both devices.

We thank you again for your interest in this experiment and for taking the time to perform the task. If you have any questions, please do not hesitate to ask the experimenter. If you should feel uncomfortable in any way throughout the experiment, please tell the experimenter. This is not a trial of how much torture you can take, but a test of the usefulness or otherwise of two experimental tools designed to help visually disabled people.

B. Instructions for Use of the Keyboard

With the "key"-device in front of you, you get access to a page of information, displayed by the computer. The page has a total of 40 lines. You have an imaginary pointer which can be moved around on that page. You can hear a word spoken when the pointer is pointing at the word or series of words. You are provided with a total of 14 different keys, which allow you to move the pointer to any location on that page and hear the word or words which are displayed in that location spoken for you.

The keys can be grouped into character-, word-, and line-keys, as well as some other keys. Each key will move the pointer to some location, and then speak the information there, unless you release it right away. If you do release the key right away, the pointer will only move, and "beep" in a high frequency tone, to let you know that the movement occurred. At this point, you should probably hit one of the keys to find out what that sounds like.

The next thing to know is that any key is able to "tell" you what it does rather than doing it. To use this feature, you hold down the shift-key while striking the key you wish to know about. So, if you forget what a key does, this will help you to find out again.

Now let's turn to the detailed key functions: Let us begin with the "current word" key, the one with the mark on it. The experimenter will help you to find this, or any other key, during training. The "current word" key will move the pointer to the beginning of the current word, and then speak it, unless you release the key right away. To the right of this key, you will find the "next word" key. This will move the pointer to the next word and speak it, and if you still hold the key down, it will move on and eventually read the whole screen for you. When it reaches the end of a line, it "beeps" and then jumps to the beginning of the next line.

To the left of "current word," you will find the "previous word" key. It works somewhat like "next word," except that it moves in the reverse direction.

Below the word keys, you will find the three character keys. Again, these behave just like the word keys, just with characters at a time.

Above the word keys, you will find the three "line-keys," again like the word keys. The current line key moves you to the beginning of the line and starts reading it. Previous and next-line move the pointer to the beginning of the appropriate line, and start reading the new line. However, you do not have to listen to the whole line. The machine will only read as long as you hold the key down. If you hold the key down long enough, it will read to the end of the line and stop there. In any case, your cursor will be pointing to the word which you heard last. Thus, "next word" will continue reading from there, "previous word" will take you back one word. Pressing "current line" again will result in starting the same line over again.

As the line keys always move you to the beginning of a line, you need two more keys, if you want to move up and down a column. So the key above "previous line" takes you exactly 1 word up, while the key above "next line" will take you exactly 1 word down. If you hold either of the keys down, it will attempt to read the current word, and then moves on to the next/previous line, to repeat same function. However, if the columns are not aligned properly, you may have to hit "previous word" in order to find the right item.

All of the keys mentioned so far move the pointer relative to where it was before. There are two "absolute" motions: To the right of the "next line" key, there is "beginning of page," which takes you to the beginning of line 1. Below that key, you find "end of page," which takes you to the beginning of line 40, which is the last line on the page.

There is one more key that might help you when you are lost: the big key in the lower left corner. It tells you the current location, by line and column, of the pointer.

You are encouraged to try all of these keys as long as you like, until you feel reasonably familiar with them. Feel free to ask the experimenter if you need help or explanations.

Please tell the experimenter when you are ready to proceed with the experiment.

C. Instructions for Use of the Speech-Pad

In front of you you find a surface, much like the page of a book, with lines embossed on it. Altogether, there are 40 lines on the page. With the little antenna attached to your finger, you can use this device to read what is written on the page-whichever word you point to will be spoken by the machine.

Thus, you can move along a line, and the line will be read to you, until you stop or take your finger away. You can also point to an arbitrary location and hear the word which is written there spoken. To repeat a word, you lift your finger and lower it down in the same place.

There are several special situations: First, if the device does not speak while you are pointing at some spot on the surface, there are probably just spaces under your finger. Secondly, if you touch a blank line, you will hear a "beep"; this will be repeated if you keep on moving on that line. Thirdly, if you reach the end of a line, so that only blank spaces follow to the right, you will hear the same "beeps," but considerably shorter. Fourthly, if you want to read backwards, you will have to move quite a bit slower than when going forwards, otherwise the device won't react. This avoids lines being read backwards accidentally when skipping backwards to the beginning of a line. This is basically all you have to know about the main area.

Special provisions are made for scanning. On the lefthand side of the page there is a specially marked column. If you move your finger into that column, at a line position, you will hear the whole line spoken, without having to trace it. The machine will stop reading the line, when you lift your finger, that is, you don't have to listen to the whole line.

If you want to access characters at a time, rather than words, you will have to switch the device into spell-mode. You do this by touching the button on the right side, in the middle of the page. To get back to word-mode, touch the same button again.

You are encouraged to try the procedures which were just described as long as you like, until you feel reasonably familiar with the device. Feel free to ask the experimenter for help or explanations.

Please tell the experimenter when you are ready to proceed with the experiment.

D. Instructions for the Target Task

The first task in this experiment can be considered a game.

First of all, each line of the page contains five evenly spaced + signs, which form columns across the lines. Overlayed with that structure, there are four different targets on the page. The layout of the page will remain unchanged throughout the task.

Each target has the same diamond shape. In the center of the target, there will be a single letter, with a star on each side of it. This basic group is surrounded by numbers. A "3" signals that you touched the outmost ring, while a "2" or "1" means that you are zeroing in to the target.

You are asked to explore the whole page, find all four targets, and remember roughly where each letter is located. When you feel comfortable with the display, you proceed to the task.

To find out what you are supposed to do, move to the top line of the page and read it. In the first stage of the task, it will, for example, ask you to "find 'r'." Now you move to the target containing the letter 'r,' move your pointer (or your finger) to the letter, and mark it by saying "mark." After eight trials of this nature, you will proceed to the second stage of the task.

In the second stage, you will be asked to spell a combination of three letters. You are asked to find the targets in th right order and mark the letters in them by saying "mark" each time you find one. If you forget what your task was, move back to the top line. It will be there for you to read it again.

Remember, to mark a letter, you move the pointer to the letter and say "mark." The machine will recognize this command. Make sure that the pointer really is pointing at the letter before you say "mark." If you are not pointing at a letter, or at a wrong letter, when you attempt to mark it, the machine will tell you so, and ask you to try again. Otherwise, the machine will answer "mark set on…" The trial is finished when you succeed in marking the last letter of the word. You will then proceed to the next trial, when you are ready, by moving to the top line again.

In all trials, work as fast and accurate as you comfortably can without making too many errors for your liking. However, you should not feel under time-stress. You may rest between the trials.

If you have any questions, feel free to ask the experimenter.

Please tell the experimenter when you are ready to start your training.

E. Instructions for the Editing Task

In this session of the experiment, you will be presented with a kind of editing task.

The display for this task will contain everything you need to perform the task. On the top of the page, there will be the current task. Near the middle of the page, you will find a sentence, which is to be changed, one word at a time. To the right on the page, there is a column of new words which can be used to change the sentence. This column is separated from the main text area by a vertical line. Near the bottom of the page, you will find the three commands insert, delete, and change. These are separated from the main text with a dashed line. The experimenter will tell you what the sentence is and what words are in the list.

When you have familiarized yourself with the display, you will proceed to the tasks. As in the target-task, you will start a trial by moving to the top line and reading the task. This task will remain there throughout the trial, in case you need to re-read it.

In the first stage, you will simply he asked to find a word on the display. This word can be in the sentence, the column of new words, or the command area. You are expected to find that word and mark it by saying "MARK," just like in the target task.

In the second stage, you will be asked to perform more complex commands, like "change cat to dog." The following is an example for going about this.

First, in all cases, you need to find the word in the sentence, "cat" in this case. Say "mark" when you have found it. The machine will then confirm saying: "mark set on cat."

Next you will have to find the word you want "cat" changed to, in this case "dog." You can find all the new words on the right side of the page, arranged in a doublespaced column. When you have found "dog" in that column, say "mark" again. The machine will again confirm.

With the two words to be exchanged marked, you need to find the command "change." The three commands are found near the bottom of the page, with dashed lines above and below them. Having found "change," you are asked to say "mark" again. In this case, the machine will answer "changing cat to dog" and exchange the two words.

Again, as in the previous task, the machine will not set a mark when you are not pointing at the right item, but will ask you to try again.

In order to provide some clues as to where to find the new word in the list at the side, the list is organized similar to the sentence, i.e. if you change a word early in the sentence, it will be near the top of the list. Similarly, if you change a word near the end of the sentence, the corresponding word will be near the end of the list .

You may ask the experimenter for help and explanations. Also, you will keep making changes to the same sentence, so eventually you will be familiar with the sentence and find the item right away.

F. Pretest Questionnaire (Completed During Session 1)

```
NAME: ________________________________        PHONE-NUMBER________
------------------------------------------------------------------
( to be filled in by the experimenter:  Subject-Number ________ )

      PRE-TEST QUESTIONNAIRE FOR TALKING TERMINAL EXPERIMENT
      ======================================================

*** ALL INFORMATION SUPPLIED BY THE SUBJECT IS STRICTLY CONFIDENTIAL ***
If you need help to understand some of these rather technical questions,
                           PLEASE ASK.
GENERAL
-------
AGE: ___   SEX:___
EDUCATION : HIGH SCHOOL _______________     GRADE COMPLETED :____
            POST-SECONDARY :___________     MAJOR :______________
            DEGREE(EXPECTED/AWARDED) ______     YEAR :_______
            OTHER :_______________________________________________
PREFERRED HAND : RIGHT :     LEFT :
(Your preferred hand is usually the one you normally write with)
******************************************************************
PREVIOUS EXPERIENCE
-------------------
a) HAVE YOU HAD ANY EXPERIENCE WITH SYNTHETIC SPEECH ? YES :__   NO :__
    If YES please give details : TALKING TERMINAL :____________(name)
                                 OTHER :  _______________________
           TYPE OF EXPERIENCE : _________________________________
b) HAVE YOU EVER USED A COMPUTER ? YES :___    NO :___
    If YES please indicate:
         Length of experience :                    ______(years).
       Average usage of a terminal/keyboard:       ______(hours/week)
       Number of courses taken in which
         a computer was used                       ______

c) HAVE YOU EVER USED A TYPEWRITER/KEYBOARD ? YES :___   NO :___
    If YES please indicate level of proficiency :
     TOUCH TYPIST :___   MODERATELY PROFICIENT :___    NOVICE :___
     (Typing pre-test :______ w/min.)
******************************************************************
```

SPECIFIC

DO YOU SUFFER FROM ANY VISUAL IMPAIRMENT? YES: ___ NO: ___

IF YES: TYPE OF IMPAIRMENT (DYSLEXIA, LOSS OF VISION...) ______________

AGE OF ONSET OF IMPAIRMENT : _____

GRADE OF RESIDUAL VISION : _____

OTHER DISABILITIES ? : ______________________________

ESTIMATE OF HEARING :

ABOVE AVERAGE : ___ AVERAGE : ___ BELOW AVERAGE : ___

EXPERIENCE WITH BRAILLE:

MUCH : ____ MODERATE : ____ LITTLE : ___ NONE: ___

BRAILLE SPEED : _______ words/min.

HOBBIES ? : ______________________________

****

THANK YOU VERY MUCH !

****

G. Post-test Questionnaire (Completed After All Experimental Sessions Finished)

POST-TEST QUESTIONNAIRE FOR TALKING TERMINAL EXPERIMENT
===

Subject-number: ______________ (to be filled in by the experimenter)

In the following questions, please place a mark on each scale to indicate the strength of feeling you have about the corresponding statement.

	strongly agree		undecided		strongly disagree
1.) The KEYBOARD is hard to learn.	|-----|-----|-----|-----|				
2.) The SPEECH-PAD is hard to learn.	|-----|-----|-----|-----|				
3.) It is easy to become confused using the SPEECH-PAD.	|-----|-----|-----|-----|				
4.) It is easy to become confused using the KEYBOARD.	|-----|-----|-----|-----|				
5.) I was usually able to complete the task more quickly using the KEYBOARD.	|-----|-----|-----|-----|				
6.) I was usually able to find the place I wanted more quickly using the SPEECH-PAD.	|-----|-----|-----|-----|				

Which device was easiest to learn? KEYBOARD: __ SPEECH-PAD: __ SAME: __

Which device caused more confusion? KEYBOARD: __ SPEECH-PAD: __ SAME: __

Which device did you prefer using? KEYBOARD: __ SPEECH-PAD: __ SAME: __

Would you recommend either device to a friend? YES: ____ NO: ____

if YES: Which one? KEYBOARD: _____ SPEECH-PAD: _____

If you could have your choice of only one device, which one would you choose?

KEYBOARD: _____ SPEECH-PAD: _____

Please give us any comments you might have about your experience in this experiment: ______________________________

____

____

APPENDIX B
SUMMARY OF EXPERIMENTAL RESULTS

TABLE B-I
INDIVIDUAL MEANS FOR ALL MEASURES IN THE TARGET PRETEST[a]

Subject	time kb	(var)	time sp	(var)	errors kb	errors sp	words kb	words sp	task kb	task sp	reversals kb	reversals sp
1	2204.5	3467455	988.5	366499	0.0	0.0	62.1	27.9	1.1	1.3	1.6	4.9
2	687.9	121976	344.3	11772	0.1	0.0	15.5	12.5	1.0	1.4	1.0	2.5
3	2230.1	3855684	975.4	239945	0.3	0.0	23.0	53.0	1.5	1.4	6.6	2.9
4	1106.3	1001002	799.5	70525	0.1	0.0	28.0	17.4	1.1	1.5	1.9	4.0
5	1302.4	956303	583.6	87418	0.1	0.1	19.1	17.8	1.3	1.1	2.8	14.5
6	680.6	99776	596.6	98410	0.0	0.0	25.5	24.0	1.0	1.1	0.4	1.1
7	816.8	378774	560.9	91981	0.1	0.0	21.6	18.3	1.4	1.1	2.6	11.5
8	1632.0	1112546	679.6	204652	0.1	0.0	23.5	16.9	1.4	1.5	4.1	9.9
9	1004.1	834371	450.4	41703	0.0	0.8	26.4	15.4	1.1	1.4	2.3	7.0
10	867.5	521762	585.1	72850	0.1	0.0	33.6	23.9	1.3	1.5	2.3	8.0
11	532.5	78747	1019.9	1164184	0.0	0.1	19.1	56.9	1.0	1.0	1.1	10.3
12	2414.9	3188342	919.5	97670	0.3	0.1	51.0	39.9	1.1	1.4	4.1	6.6
13	569.0	77687	364.6	4874	0.1	0.0	16.6	14.4	1.0	1.3	0.8	4.9
14	983.6	620033	755.4	344226	0.0	0.0	17.8	30.0	1.1	1.1	2.6	16.4
15	1266.0	821845	934.0	214115	0.3	0.0	58.0	44.1	1.1	1.5	0.9	3.6
16	1953.6	2273357	594.3	155298	0.3	0.1	70.4	32.5	1.5	1.1	8.0	4.8
17	2078.8	6924474	1084.5	201384	0.0	0.0	42.6	41.0	1.3	2.0	3.5	8.5

[a]Time is measured in 50-ms units, thus 1000 units = 50 s.

TABLE B-II
INDIVIDUAL MEANS FOR ALL MEASURES IN THE TARGET MAIN TEST[a]

Subject	time kb	(var)	time sp	(var)	errors kb	errors sp	words kb	words sp	task kb	task sp	reversals kb	reversals sp
1	2785.5	144497	2297.0	251390	0.0	0.3	66.8	54.5	1.0	2.3	7.3	20.0
2	2141.8	1019793	2047.4	699456	0.3	0.5	51.0	70.0	2.0	2.9	12.8	31.6
3	3270.0	2621543	1556.1	478489	0.3	0.3	52.6	50.9	1.4	1.9	10.1	13.8
4	2032.9	422257	1944.1	150684	0.1	0.3	56.6	41.4	2.4	2.1	4.9	12.6
5	2769.9	274471	1207.6	176656	0.3	0.0	52.4	38.6	1.1	2.9	4.0	26.0
6	1819.3	131042	1250.5	5993	0.1	0.0	61.1	55.4	1.6	1.5	3.4	3.1
7	1395.9	423541	1305.3	125929	0.3	0.0	43.3	53.0	1.5	1.6	6.9	19.0
8	2087.5	336990	1426.1	312028	0.0	0.3	39.3	41.4	1.6	1.8	6.3	15.4
9	1540.4	209788	1383.0	231812	0.1	0.1	39.8	41.4	1.9	2.5	4.8	19.3
10	1665.9	124451	1483.5	674073	0.1	0.1	70.1	50.8	1.6	1.9	6.4	20.5
11	1181.0	125679	1438.5	259633	0.1	0.3	32.1	50.0	1.9	1.4	3.8	24.5
12	2579.5	958312	2966.4	1604603	0.1	1.0	65.6	104.0	1.6	1.1	6.4	17.9
13	1421.8	117820	1054.8	147210	0.1	0.0	56.5	53.4	1.8	1.3	4.5	16.4
14	2199.3	504635	892.5	84163	0.3	0.3	50.5	35.9	1.5	1.8	10.0	10.8
15	2895.0	647457	2598.0	960241	0.4	0.4	102.9	137.0	2.5	2.9	6.6	13.8
16	3148.8	1674648	964.8	88638	0.4	0.1	102.6	54.8	3.1	1.6	16.1	15.0
17	2592.3	665914	1965.9	479277	0.1	0.3	61.3	74.0	1.1	1.1	7.0	30.9

[a]Time is measured in 50-ms units, thus 1000 units = 50 s.

TABLE B-III
INDIVIDUAL MEANS FOR ALL MEASURES IN THE EDIT PRETEST[a]

Subject	time kb	(var)	time sp	(var)	errors kb	errors sp	words kb	words sp	task kb	task sp	reversals kb	reversals sp
1	1045.2	175909	2207.5	9689337	0.3	0.0	24.3	41.5	1.0	1.2	0.8	19.0
2	897.8	209088	822.3	816548	0.3	0.0	34.5	14.8	1.2	1.0	1.3	6.8
3	1564.0	1271161	977.5	412540	0.2	0.0	49.0	27.3	1.5	1.2	2.3	5.5
4	1319.0	193087	794.5	304841	0.0	0.2	29.8	17.5	1.3	1.2	1.8	11.5
5	1305.5	619658	2244.3	2504882	0.0	0.8	36.0	52.0	1.2	1.8	2.0	25.0
6	618.8	97459	583.5	91679	0.2	0.0	27.5	23.8	0.8	1.0	0.2	1.8
7	997.8	469435	358.2	16994	0.2	0.2	36.2	9.2	1.5	1.0	2.2	5.8
8	1462.3	3905712	582.8	361226	0.0	0.0	40.5	12.0	1.2	1.2	1.7	4.8
9	830.3	141024	655.2	113172	0.2	0.0	23.0	18.5	1.2	1.0	1.0	13.0
10	641.3	90790	270.8	4620	0.5	0.2	23.3	9.8	1.0	1.0	1.0	3.3
11	686.2	29927	672.8	228454	0.0	0.3	30.0	17.3	1.0	1.5	2.0	10.3
12	1835.7	929016	883.7	147442	0.2	0.0	45.5	28.7	1.2	1.3	3.2	2.5
13	1918.5	2200370	311.5	8243	0.2	0.3	49.3	8.7	1.3	1.0	5.8	3.0
14	615.3	165411	534.2	21718	0.2	0.7	20.3	16.5	1.0	1.3	2.2	13.7
15	895.2	350051	868.2	212168	0.3	0.0	32.5	32.2	1.3	1.0	0.8	4.0
16	881.3	503742	458.5	44561	0.2	0.2	30.0	20.5	1.7	1.2	2.7	3.2
17	712.5	27476	929.0	749900	0.0	0.0	27.5	26.3	1.0	1.3	2.2	17.7

[a]Time is measured in 50-ms units, thus 1000 units = 50 s.

TABLE B-IV
INDIVIDUAL MEANS FOR ALL MEASURES IN THE EDIT MAIN TEST[a]

Subject	time kb	(var)	time sp	(var)	errors kb	errors sp	words kb	words sp	task kb	task sp	reversals kb	reversals sp
1	2466.0	1930479	1280.0	446776	0.3	0.0	56.3	28.1	1.4	1.4	6.1	7.9
2	1656.1	453968	2133.0	2527175	0.0	0.9	57.1	36.7	1.6	2.0	4.3	12.6
3	2203.0	2615212	2490.6	3614450	0.1	0.7	67.0	88.4	1.0	1.9	5.6	13.3
4	2685.1	2816259	1508.9	574248	0.1	0.3	71.0	33.7	1.9	2.0	7.6	14.6
5	2139.1	1537451	2096.6	1120443	0.1	0.6	61.7	66.7	2.1	2.0	5.6	18.6
6	1411.9	563678	966.3	268845	0.1	0.1	57.6	30.7	1.3	1.4	3.1	6.6
7	1695.0	663312	1747.4	1172711	0.0	0.3	56.4	39.9	1.7	2.0	4.0	19.6
8	2169.1	1767928	2330.7	3977871	0.1	0.4	62.1	48.4	1.4	1.3	7.9	25.4
9	1979.4	1723307	1355.3	761119	0.6	0.1	59.6	29.9	1.9	1.0	6.7	20.7
10	1450.0	270462	1001.4	116727	0.9	0.3	47.6	27.3	1.3	1.6	5.4	14.0
11	1441.4	420110	1081.1	160326	0.3	0.6	55.4	29.9	2.0	1.1	6.6	12.3
12	2795.3	3616979	2150.3	648743	0.7	0.1	76.6	65.1	1.9	1.1	8.4	5.9
13	2232.0	977621	1088.1	1300903	0.1	0.6	55.0	27.3	1.0	1.0	8.1	14.7
14	2009.1	385175	894.0	152627	0.1	0.3	57.1	14.9	1.7	1.0	9.0	11.6
15	2215.4	1013939	1860.0	1242764	0.6	0.6	69.4	69.4	2.4	2.1	6.7	9.3
16	1608.4	453456	1016.3	226177	0.3	0.1	54.3	33.3	2.1	1.3	9.3	14.3
17	2009.6	1354903	1090.4	169513	0.0	0.1	70.1	29.3	1.1	1.3	9.4	15.4

[a]Time is measured in 50-ms units, thus 1000 units = 50 s.

TABLE B-V
INDIVIDUAL POST-TEST QUESTIONNAIRES RESULTS[a]

ques. no.	answer from subject number 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
(1)	-1.5	-2	0	2	-2	1	-2	-2	-2	-1	-1	-1	-1	-1	1	-1	-1
(2)	-1.5	-2	0	-2	-1	-1	0	0	-2	1	-1	-2	-2	-1	-1	-2	-1
(3)	-2	1	-1	-1	1	-2	1	1	-2	-1	1	-1	-2	-2	-1	0	-1
(4)	-1.5	-1	1	1	-1	1	-1	-2	1	-2	-1	1	0	1	-1	1	-1
(5)	-2	1	1	-2	0	-1	2	0	0	2	1	-2	-2	-1	-1	-2	0
(6)	2	2	-1	2	0	1	-1	0	1	-0.5	-1	2	2	2	1	1	1
(7)	0	s	s	s	0	s	k	k	0	k	s	s	0	s	s	s	0
(8)	k	0	k	k	s	k	s	s	k	s	s	k	k	k	k	k	0
(9)	k	k	k	s	s	s	k	k	0	k	k	s	s	s	k	s	0
(10)	k	s	k	s	(s)	s	k	k	0	k	k	s	s	0	s	0	(s)
(11)	k	s	k	0	(s)	s	k	k	s	k	k	s	s	s	k	s	s

[a]s implies a preference for the speechpad, or some perceived advantage for the speechpad. (s) implies the same thing, if some condition could be fulfilled. k implies a preference or perceived advantage for the keypad. 0 implies that neither device was preferred or seemed to offer an advantage.

ACKNOWLEDGEMENT

The authors would like to acknowledge with gratitude the efforts of the students taking the graduate course in human-computer interaction at the University of Calgary in 1984 for their assistance in setting up and running the experiment. They would also like to thank several members of the Man-Machine Laboratory group at Calgary for comments on early drafts of this manuscript (Bruce Conrad, Richard Esau, Bruce MacDonald and Ian Witten), and especially acknowledge the considerable input in this project from Richard Esau, who has been closely involved with the practical details (software and implementation) at all stages of the project. Sheldon Maloff, who is currently rewriting the system software to a releasable standard, and is making his own contributions in the process, deserves thanks. Finally, they would like to thank two anonymous reviewers for encouragement, and for some highly pertinent constructive criticisms that helped to improve the clarity of the paper, and also detected some problems in the statistical presentation.

REFERENCES

[1] S. I. Allen, D. C. Songco, P. S. Plexico, and R. A. Morford, "A voice output module for a blind programmer," J. Visual Impairment and Blindness, vol. 75, no. 4, pp. 157-161, Apr. 1981.

[2] I. D. Benest and G. Jones, "Computer emulation of books," IEE Int. Conf. Man-Machine Syst., UMIST, Manchester, UK, 6-9 July 1982, IEE Conference Publication 212, pp. 267-271, 1981

[3] 5. K. Card, W. K. English, and B. Burr. "Evaluation of mouse, rate-controlled isometric joystick, step keys and text keys for text selection on a CRT," *Ergonomics*, vol. 21, no. 8, pp 601-613, Aug. 1978.

[4] W. J. Dixon, M. B. Brown, L. Engleman, J. W. Franc, M. A. Hill, R. t. Jennrich, and J. D. Toporek, *BMDP Statistical Software*. Berkeley, CA: Univ. of Calif. Press, 1985

[5] C. Döhrn, "The speech-pad: Direct manipulation computer access for the visually disabled based on speech and touch," *M.Sc. Thesis*, Univ. of Calgary, Dept. of Computer Sei.. Apr. 1984. ch. 3

[6] J. D. Foley and V. Wallace, "The art of natural graphic manmachine conversion," in *Proc. IEEE*, vol. 62, no. 4, pp. 462-471, 1974.

[7] D. R. Hill, C. Döhrn, J. Darragh, R. CL Esau, D. Levinson, B, Unger, and I. H. Witten, "Using speech as an output medium for computer dialogue," in *Proc. CIPS Session 84*. Calgary, 9-t t May 1984 pp. 373-380. Toronto, ON, Canada: Canadian Information Processing Society.

[8] G. A. Miller, The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, vol. 63, no. 2, pp 81-96, Mar. 1956.

[9] N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinhrenner, and D. H. Bent, *Statistical Package for the Social Sciences*, 2nd ad. New York: McGraw-Hill, 1975.

David R. Hill (M'68) earned his first degree in engineering from Cambridge University and continued post-graduate studies in ergonomics and cybernetics at Loughborough University.

He worked at Standard Telecommunication Laboratories (ITT Europe) in artificial intelligence and human-computer interaction using speech. He is currently a full professor in the Computer Science Department at the University of Calgary, where he directs the Man-Machine Systems Laboratory. He is also the co -editor of *International Journal of Man -Machine Studies*. He is a member of Sigma Xi and other professional societies.

Christiane Grieb obtained her first degree in computer engineering at Fachhochschule Wedcl, Germany, she continued her graduate studies in computer science at the University of Calgary, Canada, with her main interest directed towards human-computer interaction, especially using speech. She has now moved to Stuttgart, Germany, where she is a systems programmer developing operating system user interfaces for IBM Germany.