



JRRD

Journal of Rehabilitation Research & Development

Journal of Rehabilitation Research and
Development Vol. 36 No. 4, October 1999

Impact of digital miniaturization and networked topologies on access to next generation telecommunication by people with visual disabilities

Gregg C. Vanderheiden, PhD

Trace Research and Development Center, University of Wisconsin at Madison, Madison, WI 53705

This material is based upon work supported by the National Institute on Disability and Rehabilitation Research (NIDRR) of the Department of Education, Washington, DC 22202. The opinions presented are those of the author and do not necessarily reflect those of the Department of Education.

Address all correspondence and requests for reprints to: Gregg C. Vanderheiden, PhD, Trace Research and Development Center, University of Wisconsin at Madison, Room 360, Mechanical Engineering Building, Madison, WI 53705; email: gv@trace.wisc.edu.

Abstract —

In the past, telecommunication technologies did not present any particular problem for persons with visual disabilities. The telephones themselves were auditory in nature and could be operated by touch. As telecommunication begins to incorporate video and as telecommunication devices become more complex (including the incorporation of visual displays), new barriers are appearing. Fortunately, advancing technologies are also providing new opportunities for access.

The rapidly shrinking size and cost of electronics is allowing us to build intelligence and flexibility into telecommunication products. Advances will soon allow voice to be incorporated into most devices. In addition, clever use of networks and network-based services will allow access features to be built directly into the network, providing access to key visual information. As a result, future telecommunication systems can be more accessible than technologies of the past--if they are implemented correctly.

Key words: accessibility, disability, handicap, telecommunication, teleconference, telephone, universal design, voice.

INTRODUCTION

Traditional telephones have not posed much of a problem for people with visual disabilities. All of the features of the original dial telephones could be tactilely discerned quite easily, and, with a little practice, numbers could be dialed by touch. The greatest difficulties came from the early use of letters in phone numbers (for example, PArkway 3-3462) rather than using all numbers. If a person cannot see and cannot remember which letters were on which keys, he has a problem. When touchtone phones came on line, the new keypad was quickly learnable and usable by touch. As long as the 3×4 block of keys was found on any phone, it was fairly easy to use the phone to make calls.

Business phones presented more of a problem, especially where individuals needed to be able to see the lights across the bottom in order to use the phone. Simple light probes, however, could be used to allow individuals to detect which lines were lit, unlit, or flashing. The interactions on the phone were all auditory and thus presented no problem to people with visual impairments alone.

Over the years, however, phones have become increasingly complex. As the number of buttons increased, it became more difficult to identify all of the keys merely by touch. In some cases, picking up the receiver did not give you a dial tone until you pressed the correct buttons. Since you could not tell which lines were busy, on hold, and so forth, this could present a problem. In addition, some phones are programmable: the buttons used to select lines 1, 2, and 3 (or more) are the same buttons as those used for other functions at other times.

Access to phones has further been complicated by the introduction of visual displays that provide a variety of functions, all of which may be inaccessible to individuals with low vision or blindness. Where LCD displays are used instead of indicator lights, even light sensors that work with older phones cannot be used. For information displayed in alphanumeric form, the problems are even greater.

Finally, LCD touch panels are beginning to appear on phones. In some cases, they are appearing on standard phones where the dial or touchpad is replaced by a touchscreen in order to provide greater flexibility and additional functions. In other cases, telephone functionality is being added to other devices such as personal digital assistants or pocket computers with a touchscreen

interface. In both of these situations, the results are "telephones" or "telephone functions" that must be operated using touchscreens. As these devices are introduced into the workplace, visually impaired individuals have problems accessing technologies that may be an integral part of their job.

Additionally, we are also seeing telecommunications in general move from being entirely audio to being audiovisual in the forms of videophones and videoconferencing. This trend introduces a new type of access problem, one that is quite different in character and nature than those faced in simply operating phones today.

Advancing Technologies and Techniques

Fortunately, three areas of development are combining to provide individuals with low vision and blindness new strategies that will allow them to access next generation telecommunication products. They are:

1. Advances in digital electronics and voice technologies.
2. New cross-disability access strategies.
3. Network based services for cross-modal translation of information.

Advances in Digital Electronics and Voice Technologies

The cost of digital processing power is dropping by a factor of 10 every 5 years, and that power itself is increasing by a factor of 10 every 4 years. It is possible now to buy a video game with more computing power than Cray supercomputers had in 1985, and this trend is continuing. As a result, microcomputers are already being built into \$16 phones and \$5 greeting cards. Scientific calculators can be purchased for less than the price of the battery used to power them 5 years ago. It is possible to build increasing processing power and memory into smaller and cheaper products. This is already allowing intelligent phones to be manufactured and sold for less than the cost of the dial or push-button phones of yesterday. This also allows phones to be designed that can behave differently for different users.

Accompanying this are advances in voice technologies. Simple digitized speech (digitally recorded speech) is already used on everything from greeting cards to cheap toys for children. It is already possible to build small vocabularies of spoken words into almost any product. In the future, it will also be possible to build in synthetic speech, capable of speaking any text, for less than the cost of the cardboard box in which products ship today. When voice becomes this cheap (likely within the next 5 years), it will be a trivial problem to build voice into almost any product. Digitized speech will appear first, allowing access to products with a limited number of fixed function buttons. Later, synthetic speech will be economical; that will provide access to devices with any type of buttons or text displays. The functions of buttons as well as the contents of displays can then be spoken aloud on request, thus allowing access to persons who cannot see (or see well).

New Cross-Disability Access Strategies

These advances in voice and electronic technology will allow manufacturers to take advantage of the new cross-disability access techniques already being developed and deployed. These new cross-disability access techniques are already allowing people with a very wide range of disabilities to access standard products. These same techniques can often be implemented in products without changing the products at all, outside of adding a small amount of software to provide alternate ways of behaving. In other cases, access only requires the addition of a simple element, such as a single button or headphone jack, that also provides benefits to all users (1).

At this time, there is only one known package of such techniques, the EZ Access™ technique sets (2,3) that allow people with a wide range of disabilities (low vision, blindness, hearing impairment, deafness, physical disability, reading problems, illiteracy, and cognitive impairment) to access a single product (see [Appendix A](#) for details). These sets are fairly straightforward and flexible, allowing them to be applied across a wide range of product types. They are already helping people with vision impairments use touchscreen kiosks in airports, shopping malls, libraries, and other community locations. A series of touchscreen voting machines, accessible by individuals with vision impairment and a host of other disabilities, including reading problems, are slated to be rolled out later this year (see **Figure 1**).



Figure 1. Cross-disability-accessible voting booths, using EZ Access technologies, will be used in elections starting February 2000. Note the distinctive, diamond-shaped QuickHelp button.

A reference design for a cellular phone has been developed¹, and the techniques are also being applied to PBX phones, fax machines, and a wide range of other telecommunication devices (as well as non-telecommunication devices such as copiers, building security keypads, and so forth). **Figure 2** shows EZ Access features on these hand-held devices.



Figure 2. Hand-held communication devices with EZ Access features for persons with low vision.

Network-Based Services for Cross-Modal Translation of Information

All of the above techniques provide access to any visual information for which there are text or auditory equivalents. However, when engaged in a video teleconference, individuals with visual impairments are still going to run into problems in dealing with the visual components. For example, if one of the other speakers says, "*And as you can see we need to have a better focus on our overall distribution plan. We need to focus on this, this, and this product while letting three products over here drift for awhile.*" Unless the individual who is blind has some information as to what the speaker is pointing to, the conversation is of limited value. This problem, of course, is not dissimilar to the situation faced in live face-to-face communication. However, it is a problem that does not arise when people are on traditional audio-only telephone calls. In that situation, no one can see the speaker and all information is passed aloud.

Description on Demand

One approach to providing access to visual information via videophone (or video teleconference) is to create a network-based service that describes that information to persons with visual impairments. This would be similar to the description services provided for movies, except that it would have to be done in real time, making it somewhat more difficult to predict when there will be gaps in conversation to allow injection of the descriptions. When participating in a video teleconference call, these individuals would have the video signal routed to a remote service, where a sighted interpreter would describe aloud the graphic information. In real life, this approach is not practical, since there are no gaps in the conversation into which you can inject the descriptions. In the telecommunication environment however, this limitation imposed by reality can be overcome using time shifting and compression.

In networked telecommunications, the interpreter monitors the session, and describes graphic content when signaled by the user (the only one who would hear the description). The signaling mechanism would allow the user to indicate when he did or did not want descriptions. Alternatively, the interpreter could make the decisions and just be overridden by the user who would signal "Don't bother to describe this" or "Please describe this now." In either case, it is probable that the user would get behind in the conversation by stopping to listen to the

description. To address this, a catch-up feature, based on speech compression, could be provided. Once users were finished listening to the description, they would rejoin the conversation at the point at which they had turned to the description, listening to the discussion at an accelerated speed until they caught up. In this way, users could listen to any visual descriptions, short or long, without having to attend to the ongoing conversation at the same time.

This technique could be implemented on an "always present" or on a "present on demand" basis. The interpreter could be brought on-line at the beginning of the session to monitor everything ("always present"). This comprehensive approach might, however, be quite expensive. Most teleconference calls may be completely understandable most of the time without any descriptions. In such cases, a "Delay and Describe" feature could be provided. With this feature a person would only invoke the description service when needed. For example, a user could log onto a video teleconference and simply listen to the conversation. If something comes up during the discussion that requires description, the user would simply press a button connecting him or her to a video description "on demand" service. The video stream (from a minute or so before the button was pushed) would then be played to the interpreter, along with the user's instructions. The interpreter would then describe the requested material. When done, the interpreter could stay on or sign off at the preference of the user.

Automated Descriptions and the "Try Harder" Technique

In the future, some types of video interpretation could be done automatically by computers. For example, in the scenario above, an individual is pointing to product names on a chart. It would not be difficult for a computer to determine the location of the person's pointing indicator (arrow, hand, or pointer), and then to do optical character recognition on the words indicated and read them aloud to the user. However, if pictures of products were used instead of words, the individual may need to rely on a human being (until we are far enough into the future that truly intelligent machine vision is possible). In such cases where we have a mixture of machine interpretation and human interpretation, it may be useful to have a "try harder" function. That is, the user might punch a button that would cause the computer to try to interpret what was happening visually. If this did not work, the individual could press a try harder button, at which time a human being might be brought into the loop. Both the machine and the human may be located at a remote site and accessed via the telecommunication network in an instant, on demand.

Such network-based services are not farfetched. Such capabilities are already in existence for persons who are deaf: they may now use a videophone to call into an "interpretation on demand" center. This center listens to the conversation and provides text or video of an interpreter signing the content. A new service called CapTel™ (Ultratec, Madison, WI) has also been recently announced; this allows individuals to make phone calls where a combination of people and computers automatically translate the telephone conversation into text. This text then appears on the telephone display, while they also listen with their residual hearing.

CONCLUSION

Advances in telecommunication technologies are creating new barriers for individuals with visual impairment. These barriers are the result of both interface design and the movement toward audio-visual telecommunication. However, the continued advance of electronics and of voice and network technologies will soon provide us with the ability to reverse this trend and provide telecommunication capabilities not only more accessible, but also more functional, than anything available in the past. The technology has already reached the point where many of these capabilities are readily achievable. That is, they can be implemented with little expense or effort. Cross-disability-accessible kiosks that provide all of the above capabilities (and more for other disabilities) are already commercially deployed--the cost for adding the voice and other access being less than 1 percent of the cost of the product. Cross-disability-accessible voting booths using these technologies are scheduled to be used in elections starting February 2000. It is even technically practical to build access into small cell phones, PBX phones, and other intelligent telecommunication products. In the near future, the cost will drop even further, so that almost anything with electronic controls can be made accessible.

Perhaps one of the most interesting aspects of these technical developments is the fact that these telecommunication capabilities, particularly the network-based services, may make face-to-face interactions more accessible as well. For example, an individual in an environment where people are gesturing, pointing to things, and so on, may literally make a videophone call to themselves. That is, they would take a small video cell phone, aim it at the event taking place in front of them, and have an interpreter describe the salient parts to them. Again, the description may be done by a computer-based system or a human, depending upon the situation. An individual at a bus stop, for example, may simply aim a video telephone at the buses to have the name of the bus route, printed across the front of the bus, read aloud by a computer-based description routine that would cost very little. If this was not working for some reason (snowstorm, poor lighting, or poor aim by the user), they then might press the try harder button, and a human being would join the effort. The human might have better ability to discern partially blocked letters, to zoom in on the correct part of the bus containing the label, or to help the individual to correct the camera aim. This would work for buses, signs, menus, tickets, scribbled memos pasted on doors, or any printed information anywhere.

Thus, telecommunications products and services may not only become more accessible, but they may provide visually disabled people with powerful new tools for their general daily activities. Most of the technologies are here today, and the rest will be here soon. Whether or not they are built into the next generation of telecommunication products and services is the only open question. Hopefully, the recent Telecommunications Act, which requires that telecommunication products be accessible "when readily achievable," will bring this new access to reality as advancing technologies make all this--and more--both readily achievable and commonplace.

The talking and low-vision kiosks and voting machines are already commercially available. Speech-enabled cell phones should be commercially available within 18 to 24 mo. Annotation on demand should be available as soon as wireless communication allows for usable video transmission (which should take awhile), but annotation or description on demand (for a price) should be available over land lines (high speed internet) within 5 years.

For further information on these and related developments, see "Designing a More Usable

World--for all" section of the Trace Center Website at: <http://trace.wisc.edu/>. It contains more information on these topics as well as continually updated links to other resources available on the Web. It also has a searchable "Custom Bibliography" feature for related print articles not on the web.

¹Vanderheiden GC. Before the FCC in the matter of implementation of section 255 of the Telecommunications Act of 1996; Ex parte comments filed July 7, 1999. (*Return to [text.](#)*)

[Appendix A](#)

REFERENCES

1. Vanderheiden GC. Cross disability access to touch screen kiosks and ATMs. *Advances in Human Factors/ Ergonomics* 1997; 21A:417-20.
2. Vanderheiden GC, Law C, Kelso D. Cross-product, cross-disability interface extensions: EZ Access. *Proceedings of the 21st Annual RESNA Conference*; 1998 Jun 26--30; Minneapolis, MN. Washington, DC: RESNA Press; 1998. p. 346-351.
3. Law CM, Vanderheiden GC. EZ Access strategies for cross-disability access to kiosks, telephones, and VCRs. *Proceedings of the Technology Initiative for the Integration of Disabled and Elderly People (TIDE) Congress*; 1998 Jun 23-25, Helsinki, Finland. 1998.

Go to [TOP](#)

[Contents](#)