DOCUMENT RESUME

ED 339 150 EC 300 760

AUTHOR	Williams, John M.
TITLE	Speech Technologies. Tech Use Guide: Using Computer Technology.
INSTITUTION	Council for Exceptional Children, Reston, VA. Center for Special Education Technology.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE	Jan 90
CONTRACT	300-87-0115
NOTE	5p.; For other tech use guides, see ED 324 842-850 and EC 300 758-769.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Accessibility (for Disabled); *Artificial Speech; *Assistive Devices (for Disabled); *Communication Aids (for Disabled); *Communication Disorders; Elementary Secondary Education; Sensory Aids; *Speech Synthesizers; Technology
IDENTIFIERS	*Speech Recognition; Voice Recognition

ABSTRACT

Speech synthesis and speech recognition systems offer access to communication and information for students with communication disabilities, thus eliminating major historical barriers to learning for these students and allowing them to participate in the school environment. This guide describes two ways of producing speech synthesis: (1) by recording, analysis, storage, and play back of a human voice; and (2) by using a set of detailed pronunciation rules. Two types of speech recognition (isolated utterances and continuous speech) are also described. Examples of individuals with visual impairments, physical disabilities, and speech impairments illustrate applications of speech synthesis and speech recognition. The guide concludes with a list of communication aids and speech recognition systems, periodicals, organizational resources, and readings. (JDD)

Center for Special Education Technology

Tech Use Guide

Using Computer Technology

Speech Technologies

Speech synthesis and speech recognition systems are reshaping the lives and education of students with disabilities. Speech synthesis, also called synthetic speech, adds a "talking" dimension to products such as computers, calculators, typewriters, watches, and adding machines, allowing ease of access to them by students with visual impairments, dyslexia, and other disabilities. The world of synthetic speech is their access to communication and information, thus eliminating major, historical barriers to learning and allowing them to participate in the school environment.

Synthetic speech also serves as substitute speech for nonvocal and nonverbal people. The inability to speak has often limited the participation of nonverbal/nonvocal children in regular school programs. Talking computers, particularly laptops, are a technological breakthrough that increases the number of students with disabilities who can be educated fully. Early speech products often had metallic or echo-sounding voices that could be difficult to understand. Many of today's speech products, however, have near-human-sounding voice quality. The improved quality has led to increased sales of speech products and a proliferation of the companies producing them.

Speech Synthesis

Synthetic speech is produced in two ways: by analysis and by using rules. With synthesis by analysis, a human voice is recorded, analyzed, stored, and played back as needed. Before being incorporated into either software or hardware, the digitized voice is compressed; the computer then acts as a digitized computer tape recorder. Entire words and phrases are stored in the computer's memory and then recalled on command. The highest quality speech output can be achieved through digitalization at a sampling rate of 10,000 to 25,000 samples per second.

With synthesis by rule, better known as text-to-speech, the computer is given a set of pronunciation rules to apply as it reads text. These rules specify sounds for components of words from phonemes, the most basic distinguishable sounds in language. The rules also specify how to combine these words to produce spoken words and phrases. Additional rules regulate pitch, determine which syllables are accented, and change intonation depending on whether the sentence is a question or a statement. Synthetic speech by rule does not restrict vocabulary: The computer can synthesize any word by following the rules. However, because the sounds are artificially produced, they often sound mechanical and robot-like. Some text-to-speech products, such as Adaptive Communication Systems's RealVoice, have natural sounding speech.

Speech synthesizers are used to read aloud printed materials. For example, the Kurzweil Reading Machine scans a page and reads it aloud, using sophisticated synthesis-by-rule algorithms to produced speech. Words are clearly enunciated and spoken with almost natural rhythm and inflection. Students with learning disabilities who use this type of machine can be given textbooks augmented with speech. Instructions and screen text can be read aloud, reducing barriers created by poor reading skills. Blind students can detect errors by listening to individual letters, spoken as they are typed. Also, when finished, the entire text can be read. Blind students can likewise edit computer programming through speech synthesis code editors.

Speech synthesis is generally provided by hardware, either through one or more microchips installed in the computer or by a separate peripheral with its own power supply and speaker that can be plugged into a port on the computer. These products can either be handheld, laptop, or the size of a desktop computer. The prices for products using speech synthesizers range from $300 for a synthesizer to more than $10,000 for the hardware. Software is less expensive than the hardware.

Speech Synthesis Users

The St. Lucy Day School of Philadelphia is a school for students with visual impairments. The 22 students ranging from first to eighth grade have daily contact with a computer using artificial speech and receive at least 45 minutes of individual training weekly. The talking computers help keep the children's attention focused on their subjects; in fact, the teachers have compared it to having a second teacher in the classroom. Sister Judith Ann Moeller, principal, says, "The presence of speech output products has opened the world of information for blind users. The students have opportunities that they never had before."

Another example is Jerry Pudlik of Farmington Hills, Michigan. An eighth-grader at Warner Middle School, Pudlik has a speech impairment as a result of being hit by a car. Sitting in his wheelchair, he begins answering questions using his RealVoice, a laptop communications aid. "I have been using speech communication aids for years. They have made it possible for me to attend school. To speak to my family and friends. I could not exist here without it," says Pudlik.

A Project of The Council for Exceptional Children
Funded by The Office of Special Education and Rehabilitative Services, U.S. Department of Education

Speech Recognition

Speech recognition provides support for students who cannot use a keyboard, roll a screen up or down, exit a file, produce a spreadsheet, or perform other tasks. Today's speech recognition systems also offer enormous potential for the nonvocal/nonverbal population. After a day or two of training, speech-impaired children, and adults, are using speech recognition systems in special schools and in rehabilitation training.

Isolated Utterances

There are two types of speech recognition: isolated utterances and continuous speech. Isolated words are single words or short phrases, such as computer commands.

In an isolated-utterance system such as Dragon Dictate, when a word is spoken, the microphone translates the word into an audio signal. The speech-processing board then converts the audio signal into digital data that will be processed by the computer.

As the speech data characterizing the word enters the computer, a recognition program compares it to the acoustic word models resident in the active vocabulary. The statistical language model helps select the most likely words sent to the display.

If the word spoken is recognized and correctly highlighted in a displayed Choice Menu, the user continues the dictation by speaking the next word, which is then entered into the text of application (e.g., word processor) and the statistics for the word is updated in the Language Model. If the word spoken is not highlighted but is in the Choice Menu, the user can push an appropriate function key to select the word. As before, the word is placed in the text and the Language Model is updated. The acoustic model in the Active Vocabulary is also updated.

If the word is not in the Choice Menu, the user either types, or in the Hands-Free Choice, speaks the first letter of the word. The Language Model then selects the "most likely" words from the Active Vocabulary. These words, if any, are then displayed in the Choice Menu. Even with very large vocabularies, a typical DragonDictate user will see over 90% of their spoken words (which are in the Active Vocabulary) on the first choice menu.

Continuous Speech

For years people have dreamed of having natural, continuous voice input for their computers applications. Continuous speech is normal speech with multiple speakers and without pauses. Continuous-speech technology is knowledge-based; it makes use of phonemes, the basic elements of speech. Speech Systems, Inc., for example, has used phonemes to develop a system that recognizes continuous speech, not only words but entire sentences and phrases. It provides speaker models able to recognize speech from a variety of speakers.

The system operates on general purpose computers. Its components include a voice input peripheral (phonetic engine and generic male and female speaker models), runtime recognition module (phonetic decoder), and a developer's toolkit (syntax compiler, standard dictionary, application interface module, test and debug tools, and performance measurements tools).

The technology is matched to applications such as training systems, database access, customer service, and reporting systems. A student, for example, can create and review a report and make corrections immediately.

Once a model is loaded, the user speaks into a microphone device connected to a phonetic engine. The speech is converted into a string of phonetic codes. If called for by an application, phonetic codes may also include information on pitch, duration, and amplitude.

The phonetic decoder, a software package written in C language, runs on the host computer. It receives the phonetic code strings over an RS-232 serial line from the phonetic engine and converts them into word-string output using the resources of the syntaxes, dictionaries, coarticulation rules, and speaker model. Each phonetic string code corresponds to a spoken utterance and is processed as a unit.

Speech recognition hardware units are needed to translate voice commands to computer actions. These units can be plug-in circuit boards or external self-contained units to use with the computer. Once installed, they generally permit the user to operate off-the-shelf software through voice commands. Many units combine speech synthesis with speech recognition.

Speech Recognition Users

Matthew Carlson lives in a Los Angeles suburb. Five mornings weekly, his father Timothy places Matthew's speech recognition system in their van and transports Matthew to a private school. Matthew is 14 years old and cannot use his hands for writing, typing, or even turning the pages of a book.

In class, Matthew sits erect in his wheelchair and smiles constantly. A microphone sits on the desk near his computer. As the teacher speaks, Matthew speaks softly into his microphone; his notes appear on the screen. Later he will review them for a test. Matthew also uses his computer to write his reports, do calculus and geometry problems, take tests, and make telephone calls.

Matthew's uncle Richard, a teacher at the school, says, "Using his speech recognition system, Matt can keep up with his classmates. This system and others similar to it give my nephew and other paralyzed children the opportunity to attend regular schools and go on to college and later to get a job."

"When I take a test, the teachers let me take them using my speech recognition system in another room. It is slower than writing, but since I have never written a word in my life with my hands, this speech recognition system is marvelous for me," says Matthew.

At home, Matthew's computer is hooked up to another computer in about 10 minutes by his father. Once it is ready, Matthew can make a telephone call, turn the lights, TV, and other appliances on and off.

In Chicago, 12-year-old Mary Kessler is eating lunch in the noisy school cafeteria with her friends. Despite the noise she has a computer in front of her so she can use its speech recognition capabilities. She lost her arms in a 1985 car accident and for two years, has been using a speech recognition system in school and at home. She calls it "Old Faithful" because it has never let her down.

"I use it for writing, making telephone calls, researching, in mathematics and science classes. In my computer programming class, I am way ahead of my friends because I develop my own programs," Mary says gleefully. She learned to use the system in two days and after three months, was comfortable enough with it to bring it to school. She keeps it there all the time, since she has another system at home.

The future promises improved speech and smaller and cheaper speech synthesis and recognition products. Some day there will be a handheld, lightweight augmentative communications aid and a similar speech recognition product device with much larger memory and storage.

Products

Communication Aids

Adaptive Communication Systems, Inc., 354 Hookstown Grade Road, Clinton, PA 15026; 412-264-2288.

Artic Technologies, 55 Park Street, Suite 2, Troy, MI 48083-2753; 313-588-7370.

Automated Functions, Inc., 6424 North 28th Street, Arlington, VA 22207; 703-536-7741.

Blazie Engineering, 3660 Mill Green Road, Street, MD 21028; 301- 879-4494.

Don Johnson Developmental Equipment, Inc., P.O. Box 639, 1000 N. Rand Road, Building 115, Wauconda, IL 60084; 312-526-2682.

Henter-Joyce, 7901 4th Street, N., Suite 211, St. Petersburg, FL 33702; 813-576-5658.

HumanWare, Inc., 6140 Horseshoe Bar Road, Suite P, Loomis, CA 95650; 916-652-7296.

Innocomp, 3195 Wagon Wheel Drive, Sodon, OH 44691; 216-262-1984.

Kurzweil Computer Products, 185 Albany Street, Cambridge, MA 02139; 800-343-0311.

Luminaud, Inc., 8688 Tyler Blvd., Mentor, OH 44060; 216-255-9082.

Prentke-Romich Co., 1022 Heyl Road, Wooster, OH 44691; 216-262-1984 or 800-642-8255.

Sensory Aids Corporation, Suite 122, 205 West Grand Avenue, Bensenville, IL 60106; 312-766-3935.

Speech Enterprises, P.O. Box 7986, Houston, TX 77270; 713-461-1006.

Telesensory Systems, Inc., 455 North Bernardo Avenue, Mountain View, CA 94039-7455; 415-960-0920.

Venture Technologies, Inc., 304-134 Abbot Street, Vancouver, B.C. V6B 2K4, Canada; 604-684-9803.

Speech Recognition Systems

Artec Systems, Inc., 9104 Red Branch Road, Columbia, MD 21045.

Bartlett Systems, 11820 Parklawn Drive, Rockville, MD 20851, 301- 231-9300.

Cherry Electrical Systems, 3600 Sunset Avenue, Waukegan, IL 60087, 312-360-3500.

Consolidated Systems Group, 6421 Hazeltime Blvd., Suite 159, Excelsior, MN 55331.

Convox, Inc., 675-D Conger Street, Eugene, OR 97402, 503-342-1271.

Dragon Systems, 90 Bridge Street, Newton, MA 02158, 617-965-5200.

Hy-Tek, 1980 Rt. 30, Sugar Grove, IL 60554, 312-466-7664.

Interstate Voice Products, 1849 West Sequoia Avenue, Orange, CA 92668, 714-937-0910.

Keytronic, P.O. Box 14867, Spokane, WA 99214.

Kurzweil AI, 41 Waverly Oaks Road, Waltham, MA 02154, 617-893- 5151.

Micromint, Inc., 4 Park Avenue, Vernon, CT, 06066, 203-871-6170.

Microphonics Technology Corp., 25-37 Street, NE, Auburn, WA 98002, 206-939-2321.

MTI, Inc., NE 29th Place, Suite 245, Bellevue, WA 98007, 206-881-1789.

NEC America, Inc., 8 Old Sod Farm Road, Milville, NY 11747, 516-753-7000.

Scott Instruments Corp., 1111 Willow Springs Drive, Denton, TX 76205, 817-387-9514.

Speech Systems, Inc., 18356 Oxnard Street, Tarzana, CA 91356, 818-881-0885.

Texas Instruments, Inc., P.O. Box 2909, M/S 2004, Austin, TX, 512- 250-7111.

The Voice Connection, 17835 Skypark Circle, Suite C, Irvine, CA 92714, 714-261-2366.

Voice Industries Corp./Verbex, 10 Madison Avenue, Morrison, NJ 07960, 201-267-7507.

Voice Learning Systems, 2265 Westwood Avenue, Suite 9, Los Angeles, CA 90064, 714-261-2366.

Votan, 4487 Technology Drive, Freemont, CA 94538, 415-490-7600.

Periodicals

ASHA, American Speech-Language-Hearing Association, 10801 Rockville Pike, Rockville, MD 20852.

Assistive Technology, the official journal of RESNA. Demos Publications, Inc., 156 Fifth Avenue, Suite 1018, New York, NY 10010.

Augmentative and Alternative Communication, publication of the International Society for Augmentative and Alternative Communication. Williams and Wilkins, 428 East Preston Street, Baltimore, MD 21202.

Closing the Gap, P.O. Box 68, Henderson, MN 56044.

Communication Outlook, Artificial Language Laboratory, Michigan State University, 405 Computer Center, East Lansing, MI 48824-1042.

Resources

American Speech-Language-Hearing Association (ASHA), 10801 Rockville Pike, Rockville, MD 20852.

Artificial Language Laboratory, Michigan State University, Department of Audiology and Speech Sciences, 405 Computer Center, East Lansing, MI 48824-1042.

IBM's National Support Center for Persons with Disabilities, P.O. Box 2150, Atlanta, GA 30055: 404-988-2733.

Rehabilitation Engineering Center on Augmentative Communication, University of Delaware, Department of Computer and Information Science, Newark, DE 19711.

The United States Society for Augmentative and Alternative Communication, c/o Barkley Memorial Center, University of Nebraska, Lincoln, NE 68588.

Readings

Berliss, J. R., Borden, P. A., & Vanderheiden, G. C. (1989). *Trace resource book: Assistive technologies for communication, control, and computer access.* Madison, WI: Trace Research and Development Center.

Burkhart, L. J. (1987). *Using computers and speech synthesis to facilitate communicative interaction with young and/or severely handicapped children.* College Park, MD: Author.

Heisel-Dewert, M., & Van Der Meiracker, M. (1987). The intelligibility of synthetic speech to learning handicapped children. *Journal of Special Education Technology*, 9(1), 38-44.

Kannenberg, P., Marquardt, T. P., & Larson, J. (1988). Speech intelligibility of two voice output communication aids. *Journal of Communication Disorders*, 21(1), 11-20.

Lahm, E.A. (Ed.). (1989). *Technology with Low Incidence Populations: Promoting Access to Education and Learning.* Reston, VA: The Center for Special Education Technology.

LaRiviere, C. & Sherblom, J. (1986). Speech synthesis and speech recognition by microcomputer. In Jerry L. Northern (Eds.), *The personal computer for speech, language, and hearing professionals.* Boston: Little Brown.

This guide was developed by John M. Williams, President and Consultant of International Centers of Technology for Disabled Persons in Sterling, Virginia. John is also the editor and publisher of *SAINT* (Special and Individual Needs Technology), a monthly consumer newsletter for disabled people.

Tech Use Guides on the following topics are available from the Center upon request:

- Guide for Teachers
- Guide for Parents
- Technology for Work, Home and Leisure
- Computer Access
- Selecting Software
- Speech Technologies
- Preschool Children
- Hearing Impairments
- Computers and Writing
- Computers and Cooperative Learning
- Visual Impairments
- Learning Disabled
- Telecommunication Networks
- Augmentative and Alternative Communication
- Mildly Handicapped

This material was developed by the Center for Special Education Technology under Contract No. 300-87-0115 with the Office of Special Education Programs, U.S. Department of Education. The content, however, does not necessarily reflect the position or policy of OSEP/ED and no official endorsement of the material should be inferred.