

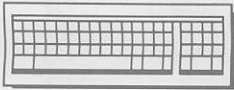
Augmentative Communication News

January, 1992 Vol 5, No.1



INSIDE THIS ISSUE . . .

Equipment



The ABC's of
Speech recognition systems

For Consumers



Applications of SR
for persons with disabilities

Clinical News



More about applications for
persons with speech impairments

University and Research



Research priorities.
Where will they lead?

Governmental



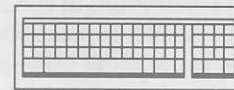
Priorities for the 1990s:
Resolve to make a difference!

Upfront

The very first issue of *ACN* four years ago highlighted speech synthesis technologies. Oh how time flies. . . Today, synthesized and digitized speech is effectively used as: 1) a prosthesis for those with speech impairments (e.g., to converse, exchange messages, use the phone); and 2) a source of information for those with impaired vision or learning disabilities (e.g., to read text, computer displays). Today, technology that allows computers to create intelligible speech, with male and female voices, is not only possible, but available to persons with disabilities. Even the general public is used to the notion "computers talk." Speech output is IN. See page 3 for more.

Talking to computers (speech input), on the other hand, is less well known, even though computers capable of recognizing human speech have been in industrial use for a decade. Sure. . . you may "talk" to your computer (*often in expletives*), but I doubt you expect it to "listen," never mind "do" what you tell it. Now, thanks to improved technology and some aggressive marketing, speech recognition is coming to the attention of the general public and the disabled community. If you haven't already heard about it, get ready! Speech recognition is the focus of this issue.

I perused the literature on speech recognition and interviewed master clinicians and researchers working in the area. Although speech recognition (*cont. on pg. 2*)



Equipment Speech recognition systems

Speech recognition (SR) is an interface between a human and a machine (computer). As such, it is like a keyboard, mouse, track ball, switch with a scanning program, or bar code reader. SR is an "intuitive" access technique, i.e., humans naturally use speech to control their environment. Other access options (e.g., typing, Morse code, or scanning) require users to learn complicated new skills. Automatic SR can improve both speed and accuracy of input, particularly in dark or low light conditions and when a user's hands are busy doing other things. Commercial applications include:

Control of machinery; Aircraft, helicopter, spacecraft control (inside vehicle); Air traffic control; Running computer networks, displays and graphics; Command and control of weapon systems; Control of toys/games; Automated training; Management of telephone dialing and calls; Medical record keeping, during or after procedures (e.g., operating rooms); Home appliance control; Data entry, e.g. inventory control; sorting packages, boxes, luggage; dictating reports, etc.

When people have impairments affecting their hands or vision, SR can enable them to do certain tasks more efficiently and with less effort. However, whether SR is the best solution, or even a good one, depends on the interaction of many variables: features of the technology, nature of the task, abilities and expectations of those involved, and adequacy of a user's training.

Basics of Speech Recognition

Table I on page 3 lists major SR companies, information about products, and examples of applications. These companies stay in business by meeting the needs of their business and industrial partners. Most are interested in how their products may benefit people (*cont. on page 2*)



UPFRONT (continued from page 1) systems are designed for able-bodied speakers, applications also can assist people with disabilities, even those with speech impairments to communicate (*write, take notes, use the phone, draw, and so on*). The **Equipment** section gives some basic information about speech recognition technology and products. **For Consumers** highlights current applications of speech recognition systems with disabled populations while **Clinical News** gives more specific information about those with speech impairments. **University/Research** and **Governmental** sections list AAC priorities for the 1990s. After all, January is a time for re-assessment and renewed resolve! Many thanks to all who contributed ideas (see list of Resources on page 8).

Finally, to everyone who completed the 1991 CEU examination (*and returned it*)... You passed! Responses to questions were comprehensive and thoughtful; I enjoyed reading each exam. Send us \$4 if you want to sign up for 1992 CEUs. ACN staff wish you a very happy 1992!

Sarah Blackstone, Ph.D., Author



Equipment (cont. from page 1)

with disabilities. But, the support for SR applications with disabled persons must come from professionals who are knowledgeable and/or companies who develop specific applications. Here are some basics:

Speaker Dependent vs. Speaker Independent: Speaker dependent means the system is user specific, and users must train it to recognize their voice patterns. Speaker independent means the SR system is designed to recognize target words, no matter who says them.

Discrete vs. Continuous: Discrete SR systems recognize words or short phrases separated by brief pauses. Continuous systems recognize connected speech.

Vocabulary: Vocabulary refers to utterances (words, phrases, sounds) the user says into a microphone to get the computer to do something. For example, if I say "hello," the computer might type *h e l l o* + space. Or, if I say "write letter," the computer might load **Word Perfect**, and bring up a template with today's date. If I say "ugh, ugh," the system may turn on lights or say a sentence. SR vocabularies can be a few utterances (e.g., yes/no) up to

thousands of words. The vocabulary size a user needs is determined by the number of commands required to control the application and whether speech will be the only access mode. Most applications don't require large vocabularies. Exception? Word processing.

Note: SR systems allow users to construct subsets or branches of vocabulary so users have access to multiple "active vocabularies" that can be activated at different times.

Microphony. A variety of microphones are available. Mics mounted on a head band or on a neck brace are preferable. Mics should be placed at the corner of the mouth, not in front. The Shure Head mounted SM 10-12 microphone, which is noise canceling, is often provided with SR systems. *Throat* mics and *in the ear* mics will improve for future use.

Deciding to use SR

SR can be a powerful tool for some people, under some circumstances. For others, it may gather dust on the shelf. When professionals/families/vendors get excited about a technology like SR, they sometimes recommend it only to discover later on the needs of the user are not met or the user was never adequately trained. These "mistakes" often give technology a "bum rap" even though the prob-

lems are not technological. To avoid recommending a SR system for someone's shelf, ask and answer the following 20 questions¹:

1. What does the person want to do? Use a computer, control equipment?
2. What size vocabulary is needed?
3. What are other available input options? Could voice input make the task easier? more efficient? more productive? Why?
4. Will voice be cost effective?
5. How often will the user use SR? frequently or seldom?
6. What kind of feedback does the user require: system action, recognition accuracy, rejects? Is it available?
7. Who will train the user?
8. How much time is available to train user?
9. Can the user choose his/her own vocabulary? If not, who will decide on the vocabulary?
10. How much time can the user spend training the vocabulary?
11. Will the vocabulary work in real situations?
12. Is tree branching of vocabulary necessary?
13. Can the user modify the vocabulary later if he/she has confusion between words?
14. Will the user turn the SR system on/off. How?
15. Can the user wear a microphone? If so, where can it be placed? What kind of mic?
16. Is sufficient space available to accommodate the equipment?
17. Does equipment need to be portable or movable? Does the user need to move while using it? Needs a wireless system?
18. Is the user under stress during normal operations? non-normal operating conditions?
19. Is the stress psychological/physiological?
20. What percentage of error is tolerated by user?

Adapted from Poock's "Things to ask yourself when deciding on a speech recognizer."

In summary, professionals should:

- 1. Be familiar with the range of available SR technologies. For starters write each company listed in Table I for information!
- 2. Know the answer to each of the 20 questions before deciding to use SR.
- 3. Select the SR system that best fits the specifications (i.e., needs and capabilities of the individual and the task.)

Then, the fun starts: Training the individual to use SR and evaluating the extent to which the SR meets the individual's needs!

AUGMENTATIVE COMMUNICATION NEWS is an independent, non-membership publication. Copyright 1992 by Augmentative Communication, Inc. Reproduce only with written consent. Author/Editor, Sarah W. Blackstone, Ph.D.; Published by Augmentative Communication, Inc., One Surf Way, Suite #215, Monterey, CA 93940

Subscriptions (per year): By personal check U.S. & Canada = \$41 U.S.; Overseas = \$52 U.S. Institutions, libraries, schools, hospitals, etc.: U.S. & Canada = \$63 U.S.; Overseas = \$74 U.S. Single issues \$10.

Note: Special rates for consumers and full-time students available.

ISSN #0897-9278

Telephone: Voice and FAX same number (408) 649-3050



Table 1. A Sampling of Speech Recognition Companies (Compiled by Gary K. Poock)

Company	Product	Type	Active Vocab	Hardware	Applications	Approx Cost	Contact
Articulate Systems	Voice Navigator II	Discrete Speaker depend.	200-1,000	MacIntosh Plus 2 MB RAM, HD	40 application setups available. With voice record software can record sounds. Not created as a dictation device	\$649-899 from dealers	600 W. Cummings Park, Woburn, MA 01801 (800) 443-7077
Dragon Systems	1. Dragon Dictate. 2. Dragon Writer	Discrete Speaker depend. adaptive	30,000	IBM 80386 8MB RAM, 40 MB HD	Has built in commands for some applications. 80,000 backup dictionary. Word boundary can be adjusted. Can teach it to ignore sounds. To offer 5 languages.	\$9,000	90 Bridge Street Newton, MA 02158 (800) TALKTYP
			1,000	2MB RAM (same)		\$2,495	
IBM	VoiceType	Discrete Speaker depend.	7000 (5,000 base/2,000 for user)	6MB RAM, 30MB HD, DOS 5.0., IBM M-Audio Capture/ Playback	DragonDictate technology withback up dictionary. Neil Squire Found. using in automated work station (4381 Gallant, N. Vancouver, B.C. V7G 1Z1 Canada)	\$3,555 (w/IBM AC & P Board)	90 Bridge Street Newton, MA 02158 (800) TALKTYP
Kurzweil	Voice Type	Discrete Speaker. Indep.	20,000	IBM 386, 16 MB RAM	Large vocabulary, speaker independent Application focus in medicine (e.g., emergency room, radiology, pathology)	\$30,000	411 Waverly Oaks Rd. Waltham, MA (617) 893-5151
Scott Instruments	SIR 20	Discrete Speaker Indep.	75 active/ 200 total	PC- XT, AT	Licenses out their technology. Some applications for handicapped (e.g., Verbotics, Inc. (214) 235-5682.	\$1,495	1111 Willow Springs Denton, TX 76205 (817)387-9514
Speech Systems, Inc.,	DS 200	Contin. Sp.Ind (semi)	3,000	Self contained work stations	Phonemic recognizer. Medical applications. Works in Unix environment	\$35,000	18366 Oxnard St Tarzana, CA 91356 (818) 881-0885
Verbex	Verbex 6000	Contin Speaker depend.	300-600	Standalone or PC card	Applications Express in NJ uses word predictor with the Verbex 6,000 (908) 389-3366.	\$6,500	1090 King Georges Post Rd., Edison, NJ 08837 (908)225-5225
Voice Connexion	Hal B	Discrete Speaker depend.	500 approx.	PC card	Environmental control. Modem control. Pop-swap/pop-dial features.	\$700-1400	17835 Skypark Cir #C Irvine, CA 92714 (714) 261-2366
VOTAN	Voice Key	Contin Speaker depend	125	PC-XT, AT card	CAD applications, etc. In 1 year VOTAN for 3.0 Windows avail. Contact J. Burnett, Wash. St. Univ. School of Architecture (509) 335-1937.	\$1,800	4487 Technology Dr Fremont, CA 94538 (510) 490-7600

REAL PROGRESS WITH SPEECH OUTPUT

No more unintelligible, robotic sounding, male-only voices! As we begin 1992, intelligible, acceptable, sex and age appropriate, computer generated speech is an affordable option. Important things are occurring.^{2,4}

Synthesized speech. The technology improved, of course, and most communication aid manufacturers now offer at least moderately intelligible speech. Also, the Infovox now has 8 languages (American & British English, Swedish, Italian, German, French, Spanish, Norwegian.)⁵ The most significant change, however, is that a major, international corporation, *with a commitment*, decided to make something happen for people with disabilities! The Digital Equipment Corporation formed the Assistive Technology Center (ATC) to make DECTalk available to persons with disabilities. For years, research has shown DECTalk to be more intelligible, more natural sounding and preferable to other synthesizers. But, DECTalk also was expensive and not very portable (i.e., heavy/not battery operated). The collaborative efforts of the ATC have made DECTalk available in 3 formats.⁶

The MultiVoice™ a small (2.6 pound), battery powered, upgraded version of DECTalk. Developed as a joint research and development project with Children's Hospital in Boston. Documentation allows users to "tweak" the system. A future version will be even smaller with an internal speaker. Available from Institute on Applied Tech-

nology, Children's Hospital, Fegan Plaza, 300 Longwood Ave. Boston, MA 02115. Also Words +, Inc. (800) 869-8521.

License agreements. Companies developing products for people with disabilities may now license DECTalk technology from the Digital Equipment Corporation. To date, agreements exist with Kurzweil, Sentient Systems, Prentke Romich Co., Phonic Ear. To inquire, contact (800) DIGITAL

DECTalk PC. Designed with features to meet the needs of persons with visual impairment. This PC version of DECTalk includes a P.C. board with speaker, floppy disks to download software, manuals on both audio tape and in ASCII, as well as print and Braille, floppy disks with Braille labels. Call (800) DIGITAL

Digitized (or recorded) speech. Having digitized speech in communication devices has made a major difference. Speech and other digitized sounds are intelligible when recorded carefully; reflect age, sex, ethnic, and linguistic characteristics of any speaker; are affordable; can meet the needs of the young and/or persons with limited communication needs and linguistic capabilities, and are easily programmed. Available devices? Macaw and Parrot (Zygo Industries); IntroTalker (Prentke Romich) Digital Augmentative Communicator (Adaptive Comm. Systems).

The Future

We've come a long way. But, speech output issues are not resolved, particularly for AAC users who depend on personal communication devices. Areas of current research are mentioned on page 6.



For Consumers

Applications for people
with disabilities

Communication devices are designed for people with communication impairments. Speech recognition (SR) products, on the other hand, are designed for able-bodied persons who have "normal" speech and intact cognition. SR is, however, a viable control site for some people with disabilities and can be preferable to scanning, light pointing, mouth or head sticks, morse code and key guard options. These methods can be extremely slow and often require the use of acceleration techniques (e.g., abbreviation-expansion, semantic encoding, prediction), which place cognitive processing demands on users we have only just begun to examine.

Many want to use SR exactly the same way as able-bodied speakers. Others, however, have needs that are more specific (e.g., controlling a robotic arm) or require adaptations (e.g., combining SR and scanning to use a word processor). Disadvantages of SR for persons with disabilities are: the cost of equipment and the training required to use it. Also, professionals and users complain about the lack of portability, difficulties mounting and stabilizing microphones, and incompatibility problems with other input methods. Populations currently benefitting from SR are people with:

High spinal cord injuries. This population is probably best matched to the technology. Current users do word processing, computer aided design (CAD), use spread sheets, write software, dial the phone, turn on/off appliances, and control robotic arms using speech. Even those on respirators can learn to time their utterances so systems recognize speech despite white noise.

Amputees, repetitive motion injuries, multiple sclerosis, etc. affecting keyboard access. While SR is sometimes appropriate, other solutions are often more cost effective. For example, ergonomic modifications to a work station and exercise,

limb prosthesis, splints, Free-wheel,TM Headmaster,TM and multi-modal inputs should be considered.

Severe visual impairments. This group requires speech output for feedback. Navigating through environments using graphic user interfaces (GUIs) is still problematic. For those with low vision, many SR system work with screen enlargers.

Severe developmental disabilities. SR may encourage users to interact with their environment and may increase vocal behaviors:

- A severely physically and mentally disabled woman learned to operate a speech recognizer to control a fan, vibrator VCR.⁷
- Subjects (ages 7-21) without functional speech who understood picture/object associations and cause/effect used the Sound-to-Speech/Graphics System to select (directly or by scanning) photographic quality images to generate a spoken message or an environmental effect.⁸

Speech impairments. Individuals with mild speech problems can use some SR systems with minimal difficulty. For those with moderate/severe dysarthrias, verbal apraxias, and other impairments, however, research shows SR is doable, but it is a huge time investment, requiring specific expertise. Studies show:

- A discrete, speaker dependent SR system may decode impaired speech better than human listeners. Eighty-one deaf, cerebral palsied and otherwise speech-impaired subjects read a 300 word vocabulary 6x over several sessions and then read a 65-word story. Results show a wide range of intelligibility. Of interest was the SR system recognized 25% of the speakers better than human listeners. Poor speakers were not understood by either. Dysarthric speakers were more difficult for the machine to recognize than the speech of deaf and hearing impaired because of their non-speech noises, limited repertoires, and inconsistent error patterns.⁹
- An adult with cerebral palsy and moderate dysarthria used the Keytronics Model 5152 V and a 28 word functional vocabulary to run an expert system and modem with 70% accuracy. Considerable time went into training and "tweaking" the system. The user felt recognition rates were "good enough" for some applications with limited vocabularies, but training was tiresome.¹⁰
- Examining the use of SR with young speech impaired children, researchers concluded SR technology was more suitable for persons with acquired speech impairments. The SR system did not accommodate the degree of variability occurring with children. Researchers also felt that SR was counterproductive to the development of communication skills even though they could increase accuracy by limiting vocabulary and choosing words produced consistently by the child.¹¹
- Studies of several individuals with cerebral palsy using two different SR systems led researchers to suggest that future studies examine factors such as voice stress, fatigue, boredom.¹²
- Comparing recognition rates of 10 dysarthric speakers with CP to 13 nondisabled speakers (ages 20-53) using the Shadow

VET 2 SR system, researchers reported accuracy rates of impaired speakers were lower, but the overall pattern of correct recognitions was similar (i.e., words better than vowels; vowels better than consonants). They concluded improvements in SR technology will benefit both groups. However, algorithms that filter out sounds beyond a certain length and frequency and/or that can compensate for inconsistencies are needed.¹³

Multiple Inputs. Instead of trying to accelerate input using a single access technique, researchers are investigating the possibility that for impaired speakers, combining multiple techniques may be effective.

- Scan and Speak is being developed in Toronto using the Voice Navigator and a scanning interface developed by project staff using HyperCard. To date, eight individuals, ages 5 to 28, with limited literacy skills who use a variety of augmentative techniques (speech is inadequate to meet daily communication) have participated. All make 3 or more repeatable utterances and can use word categories to scan. Preliminary findings show overall rates on copying tasks improve when speech recognition is added to scanning input. Cognitive factors are being considered. Subjects prefer the use of combined inputs, and knew they were faster using both.^{14,15}
- Can speech input be used by individuals with dysarthric speech as an auxiliary user interface to accomplish work-related tasks such as word processing, spread sheets, and computer aided design? The goal of Phase I (Speech Assessment) is to identify functional vocabulary sets and construct a word list of 176 words, which includes sets of monosyllabic and bi-syllabic words. 10 tokens are being collected for each word across 5 sessions. Then, words are indexed according to error matrices and order in session to determine if the SR system (the Dragon Writer) maintains models within a set and the variability of test/retest over time. Staff hope to identify subsets of words that work together to generate vocabulary lists. Phase II (Integrated motor and speech assessment) will look at keyboard, single switch and distributed switches; and also address mental load issues.^{16,17}

SR as the front end of a translation system. Maybe someday. . .

- People who do not speak well may use a computer to translate their unintelligible utterances into intelligible speech.
- SR Systems may be worn.
- SR systems may translate our words into another language when we visit a foreign country.
- People with severe hearing impairments may wear SR glasses to recognize their partner's speech and display it in text.

These "somedays" require speech recognition and synthesis technologies be on a single integrated chip (ASIC) and large vocabulary, speaker independent, continuous SR systems. For now, then, SR is one of several means of access and may even encourage some individuals to use speech in more productive and meaningful ways.



Clinical News

Use of speech recognition with impaired speakers

How to train able-bodied people to use SR systems has been studied since the 1970s. Some of that information can (and should) be applied to the use of SR with individuals who have disabilities, including those whose speech is characterized by misarticulations, inconsistencies in volume, rate, and inflection patterns, pitch breaks, and intermittent production of non-speech sounds. Our research base and clinical experiences are still very limited. However, this section offers some preliminary guidelines for consideration.^{15, 17, 18, 19}

General Guidelines

- Avoid generalizations about impaired speech. Remember there are many kinds of dysarthrias, other speech impairments, and degrees of severity.
- If the application is for environmental control, consider a discrete, low cost system
- If the application is word processing and speech is impaired, consider SR as one of multiple inputs (e.g., a way to do macros).
- Use small vocabularies. Most research suggests people with moderate/severe dysarthrias can use very limited vocabularies, often far less than 30 utterances.
- Use vocabularies with subsets.
- Be sure the microphone is placed in the optimal position and stays there.
- Tweak the vocabulary based on characteristics of the system and the speaker's patterns.
- Provide as much information to the system as possible (i.e., use multi-syllabic words/short phrases).
- Select and train systems to accomplish a specific task.
- Be sure initial reference templates are trained well.
- Present training and applications in ways that are motivating. Be aware that training and using SR systems can be very frustrating. When a SR system rejects an utterance, some interpret it as their failure.
- Be realistic and reasonable about recognition rates.

Training Considerations

We do not have protocols to train individuals with speech impairments to use SR equipment. However, protocols do exist for other populations.

- 1. Tyler Carpenter¹⁸ has trained more than 25 people with disabilities (*some with mild speech impairments*) to use SR systems and written a protocol.

Step 1. Examination of speech as a computer entry method: He notes that of those who come in wanting SR, more efficient input methods are found in about 50 percent of the cases.

Step 2. Enrollment and training: Includes developing and testing the application vocabulary.

Step 3. Instruction in the use of the application and the computer.

Approximately 14 to 20 hours is needed to set up the equipment and application program. By the end of training clients can customize their system, build their own macros and are increasing their speed. He

recommends several short sessions during the first week with additional sessions provided 2 hours per week over a series of weeks. This schedule spreads out training and allows the user to practice and explore. For more, contact him c/o the Institute on Applied Technology, Children's Hospital, 300 Longwood Avenue, Fegan Plaza, Boston, MA 92115.

- 2. Gary Poock (also ACN's publisher!) makes suggestions below. His research in training people to use SR systems for various applications is widely known.

Know How Your Speech Recognizer Works¹⁹

- #1 Does it want me to talk slow or fast? Some systems work best when you use a normal rate; others, when trained faster.
- #2 Does it have a visible gain control or volume meter? Unfortunately few systems have visible or audible feedback. Treviranus noted users seem to benefit from both and that timing of feedback is also important. When combining speech and scanning, she found user's needed the names of category echoed after it was selected. Then another auditory signal when a word is selected, then a pause and the selected word echoed.
- #3 Does it want me to whisper sweet nothings to it, or speak up like I am in charge of this affair? Encourage users to speak authoritatively. Recognizers do better.
- #4 Use the microphone in the correct manner. Find out exactly where it should be? Don't remove the foam rubber material sometimes covering the side you speak into. It's there to dampen the effects of air flow noises.
- #5 Should the user speak soft, silent and gentle at the end of words or pop out hard sounds? You'll need to test this out. Obviously some speakers have little control. You want to develop templates that recognize the utterances despite their production variations.
- #6 Does the SR system want the user to talk very consistently or vary speech during training. Systems that require speakers to be consistent when forming templates won't work for most dysarthric speakers.
- #7 Test your SR's capability to quasi word-spot, or pick out subtle differences. It may be natural for users to say "Turn it on" and "Turn it off" but SR systems may have difficulty discriminating between them. You might try "Turn it on" and "Off now."
- #8 Test its sensitivity to noise. If its a problem, try to block noise out at the site of the SR system, rather than at the noise source. Some use a court stenographer's mask and mount the SR microphone inside the mask.
- #9 Know how the system works best and teach the user all you know. For example, some SR systems work better if you use varying cadences when developing the templates.
- #10 If you are really serious about high quality results, sit with the user until the system is trained. Computer aided instruction is not as efficient and can be disastrous.
- #11 During token input and formation of reference templates in speaker dependent systems, allow no mistakes. Again, this is difficult. It requires a knowledgeable trainer to be present. Sloppy work upfront, leads to...you guessed it!
- #12 After you're all done with training, test the entire vocabulary 3 times, entirely through all the utterances. Be random and record how well each is recognized. If any utterance is not recognized 2 out of 3 times, throw out the template and retrain the utterance. Then test it again.
- #13 Don't begin serious operational use of the SR without having it all "tweaked and tuned." Do people race cars without making sure they are tweaked and tuned properly? No. Don't do it with SR systems either.

Note: The American Voice Input/Output Society (AVIOS) will keep you informed about SR applications. For information, write AVIOS, 4010 Moorpark Avenue, Suite 105K, San Jose, CA 95117.

A Note on Speech Improvement

SR systems may encourage speech production if tasks are motivating and the user's utterances are often recognized. SR technology is not, however, designed to be an articulation or voice trainer. In fact, as a client changes (i.e., improves), the SR system is less likely to reinforce/recognize the utterance. You never know. Someday, researchers may develop specific applications to improve speech production, using "smart" models working concurrently. The SR system would adapt as the user changes; and using a normative model, reward changes most closely approximating normal speech.¹⁵



University/Research

Research priorities: 1992 and beyond

Note: My sincere thanks to those contributing their ideas and visions to the University/Research and Governmental sections. See list of resources on page 8.

THE QUESTION? In your opinion, what should be the top AAC research priorities over the next few years?

THE ANSWERS. Most expressed a need for AAC to develop a stronger, more solid scientific base drawing upon knowledge from within and outside of AAC (e.g., cognitive science, ergonomics). I have listed other priorities within some general categories:

Psycho-social aspects of AAC.

1. Investigating why listeners outside the AAC/Special Ed community are put off by an AAC speaker? Is it the technology, speaker or a combination of the two? What can AAC users do to compensate?
2. Delineating social skills and strategies to maximize the AAC user's communicative opportunities.
3. Determining what applications communication aids have in real life (classrooms, workplace).
4. Investigating why some people use technology, and others don't.
5. Determining how to transfer essential information about AAC to the user's ever changing partners.
6. Focusing on issues related to peer interaction.

Decreasing the cognitive load of AAC systems.

1. Studying motor habituation as it relates to use of aids/techniques
2. Finding ways for computers to take on more of the cognitive load
3. Exploring the learning requirements of various symbol sets/systems, encoding and other techniques.
4. Incorporating human factors research in the design of communication aids and access techniques.

Interface issues

1. Developing a cognitively transparent interface (taking advantage of neural control).
2. Determining ways to construct text that take less effort, i.e., allow users to choose rather than input.
3. Developing integrated work stations allowing use of multi-media.
4. Exploring ways to translate complex thoughts into single output patterns.

Language, learning and communication issues

1. Continuing to define communicative competence. What should goals be, and what technologies can be used to reach goals?
2. Determining how to better evaluate communication skills.
3. Investigating how children acquire natural language (both comprehension and expression) using AAC techniques and using that knowledge in our early intervention programs.
4. Asking questions about how dysarthric children develop expressive language. Is their a dysarthric equivalent of the "babbling" stage of a child's in language development?
5. Determining what helps AAC users develop literacy skills & why.
6. Investigating orthographic-based systems designed for users with language/learning disabilities.
7. Investigating issues related to cognition, communication, and types of AAC interventions.

Speech (and other) technologies

1. Investigating the impact of speech output technology on the development of literacy skills.
2. Determining how listeners process the language produced by synthesized speech in discourse, in noisy environments.
3. Continuing voice source research to better model the human voice.
4. Determining under what circumstances speech recognition is a viable option for person with impaired speech and what speech characteristics are needed for good recognition.
5. Improving the naturalness of speech synthesis (i.e., better prosodics, segmental patterns).
6. Exploring ways for individuals who are at risk for losing speech (i.e., those with amyotrophic lateral sclerosis, multiple sclerosis) to capture characteristics of their voice. Then, when they lose their speech, we will know how to generate synthetic speech that sounds like them.

7. Investigating ways for synthesizers to express emotion.
8. Increasing access to multi-lingual synthesized speech.
9. Personalizing systems so they are sensitive to gender and cultural differences in conversational style.

Instructional materials:

1. Developing basic curriculum guides that are multidisciplinary, accessible to the practitioner, address real, practical barriers and apply AAC principles and practices.
2. Determining how to adapt educational curricula for AAC users.

Research methodology:

1. Developing a Handbook of Research in AAC, including sample research environments, recommendations for specific measures to encourage the use of standard designs and instruments.
2. Replicating current findings (either fully or partially). Note: If single subject designs are the way to go, we need replication to start making stronger inferences.
3. Conducting longitudinal studies in real life situations.

Efficacy of intervention methodologies

1. Investigating variables involved in Facilitated Communication.
2. Determining how much and what kind of training results in what level of success when people use communication systems.
3. Comparing approaches (e.g., interaction effects of those using and those not using voice output.)
4. Measuring and comparing the functional impact of interventions.
5. Increasing the knowledge base within AAC in interventions with regard to specific populations, e.g., aphasia and autism.

THE QUESTION? In your opinion, what topics will AAC researchers be addressing in the year 2000?

THE ANSWERS: Those interviewed predict researchers will continue to ask many of the same questions, i.e, what is the best way to teach communicative competence? how can we solve problems with rate? Most also believe technology will continue to be ahead of what we know.

- 1. We will be exploring how best to use artificial intelligence technologies in communication systems;
- 2. We will be designing and developing systems that adapt to users, i.e., take information user is able to provide the machine (gestures, vocalizations, etc) and allowing user to express affect, content, and paralinguistic characteristics in real time.
- 3. We will still be searching for real time speech output.
- 4. We will be trying to address the needs of new populations.

Some advice? Dare to search for "What is possible and desirable rather than what is feasible and what needs to be done."²⁰

ON THE HORIZON

1. Voice Output Communication Aids: An International Project on Communication Aids for the Speech Handicapped (IPCAS).

This publication by K.

Galyas, G. Fant, and S. Hunnicutt has three sections:

- 1) An historical perspective on speech synthesis technology.
- 2) Clinical and personal use of voice output communication aids
- 3) An overview of commercially available systems that use either a synthesis or stored speech, including the type of linguistic input they allow (e.g., pictures, symbols, words, spelling, abbreviation expansion)

For information about its availability contact: The Swedish Institute for the Handicapped, Box 510, Vallingby, Sweden S-162 15.

2. Staff at the Microcomputer Applications Laboratory/Augmentative Communication Services at the Hugh MacMillan Center have two exciting projects.

- 1. **Voice Switch.** In a prototype stage, it uses a lapel microphone, and will adjust for frequency and duration parameters.
- 2. **Scan and Speak,** discussed earlier in this issue on page 4, will also be made available commercially.

For more information, write to 350 Rumsey Road, Toronto, Ontario, Canada M4G 1R8



Governmental

What should our governments do?

THE QUESTION? What could government agencies do to improve the lives of persons who use AAC?

THE ANSWERS:

1. Support professional training programs in AAC.
2. Support the development of paraprofessional training (classroom assistants, community-based technical support personnel.)
3. Create an AAC/Assistive Technology service provider credential.
4. Create uniform, long term service delivery systems for citizens, regardless of age, kind of disability, economic status.
5. Encourage agencies to coordinate services and stop wasting time and resources "arguing."
6. Make sure assistive technology is considered part of all health-care reforms.
7. Lobby/persuade governments to provide funding for devices and services.
8. Address issues of accessibility and access in countries where laws do not yet exist
9. Convince governmental agencies and the private sector (large companies) to

encourage the active participation of AAC users.

10. Support consumer groups that address transportation/communication barriers.
11. Extend educational opportunities beyond age 22 years, i.e., adult education.
12. Provide mechanisms so that advocates (especially peers) can assist AAC users to coordinate services.

DO ONE THING IN 1992!

If we all did just one thing in 1992, it would make a difference! Here are some ideas.

1. Learn how to lobby governments. Consider carefully who you vote for. Pay particular attention to how local and national politicians deal with health care reforms.
2. Convince legislators, and other leaders. "Communication is an essential human right."
3. Organize letter writing campaigns to address specific, local/national issues.
4. Lobby specific governments to invest in the development of synthesized speech (e.g., Hebrew and Arabic languages.)
5. Evaluate what school districts are doing to satisfy requirements such as "least restrictive environment." Then, be prepared to do something about it.
6. Make sure AAC users have time to practice reading, writing, communicating, etc. Studies confirm the importance of automaticity.
7. Increase dialogue with consumers, and encourage them to be more demanding, particularly for adult services.

8. Parents and professionals! Take time to talk with and read to AAC users.
9. Give parents/caregivers/consumers information they need to become advocates.
10. Make opportunities for AAC users to engage in meaningful experiences.
11. Support mechanisms that protect consumers and minimize abuse of limited funding resources.
12. Increase public awareness. (e.g., hold an AAC open house, send a story to your local newspaper/T.V. or radio station) and change misperceptions.
13. Persevere with at least one AAC user until they learn to use their system.
14. Do case studies detailing what actually happens in functional situations.
15. Encourage everyone to support ISAAC and its chapters. Subscribe to AAC Communicating Together, and Communication Outlook. Attend the 1992 Biennial Conference in Philadelphia!
16. Make a professional commitment to share and keep-up-to date. Work cooperatively. Identify gaps in service & fill them. Identify overlaps and try to avoid them.
17. Subscribe to ACN. And, if you do car pool with your children, keep ACN in the car. Read it while you wait for them. You'll be a happier driver.

Note: I swear I did not pay or cajole anyone to say anything... but it is an excellent idea! S.B.

Speech I/O for Persons with Disabilities: Priorities for the 1990s

In October, a group of experts attended the CSUN Voice I/O and Persons with Disabilities Conference in Palm Springs, California. Conference goals were to establish priorities for the 1990s in speech recognition and synthesized/digitized speech for persons with disabilities. The group arrived at a consensus of the top 8 priorities for the 1990s in speech I/O for people with disabilities, after listing over 100 priority areas. Participants discussed each priority area in detail and recommended plans of action that, in their judgement, could result in the priority being accomplished. A second meeting in March at the CSUN Conference is planned to review and discuss the Speech I/O priorities listed below. For further information about the March, 1992 Conference or to order a copy of the Proceedings from the October, 1991 Conference, contact Dr. Harry Murphy, Director, Disabled Student Services, California State University-Northridge, 18111 Nordhoff Street-DVSS, Northridge, CA 91330. Phone (818) 885-2578. FAX (818) 885-4929.

RANK	PRIORITY AREA	PRIORITY STATEMENT
# 1	ACCESS TO TECHNOLOGY BY PERSONS WITH DISABILITIES	Design, evaluate, produce and distribute a range of speech I/O devices for individuals with a range of needs, abilities and economic resources at user affordable (low) costs.
# 2	STANDARDS FOR SPEECH I/O TECHNOLOGIES	Develop standards for using speech I/O, that address the human/machine interface, the hardware, and the software.
# 3	COLLABORATION IN THE DEVELOPMENT OF SPEECH I/O PRODUCTS	Increase interdisciplinary and vertical collaboration among the end user, primary caregiver, service provider, engineers (hardware/software), linguist, phonetician, human factors specialist and industry management to bring about the development of appropriate products. (see #4)
# 4	DESCRIPTION OF CURRENT AND POTENTIAL MARKETS	Identify present and potential consumers by disability and application needs.
# 5	INTUITIVE, FUNCTIONAL INTERFACES	Advance the development of intuitive, functional interfaces able to facilitate the human/machine communication/control process.
# 6	PERSONALIZED/MULTILINGUAL SYNTHESIS	Develop voice output systems that enable articulation of all human sounds, with speech qualities and characteristics appropriate to or requested by the user.
# 7	TELEPHONE AS A SPEECH I/O DEVICE	Leverage aspects of portability, policy and intelligent networks to enhance the telephone as an I/O device for people with disabilities.
# 8	A COMPREHENSIVE, GOVERNMENT SPONSORED, LONG RANGE PROGRAM	Create a structured national program with a long term charter for research, development and implementation of voice I/O for people with disabilities, which has well defined objectives, milestones, measurable outcomes and coordinates academic researchers, sub-contractor companies, clinicians and consumers.



REFERENCES

- ¹ Pooch, G. (1985). Things to ask yourself when trying to decide on a speech recognizer. Unpublished checklist.
- ² Edwards, A. (1991). *Speech Synthesis: Technology for Disabled People*. Baltimore, MD: Paul H. Brookes Publishing Co., Inc.
- ³ Murray, I.R., Arnott, J., Alm, N., & Newell, A. (1991). Emotional synthetic speech in an integrated communication prosthesis. *Proceedings of the 14th Annual Conference*. Resna Press: 311-313.
- ⁴ Higginbotham, J. & Drazek, A. (December, 1991). *Speech Output. Communicating Together*. 9:4, 11-13.
- ⁵ Sheri Hunnicutt. (December, 1991). Personal communication.
- ⁶ Tony Vitale. (December, 1991). Personal communication.
- ⁷ Brown, C. & Cavalier, A. (1986). Persons with profound mental retardation speak out and take control. *Proceedings of the American Voice Input/Output Society (AVIOS)*. San Jose: AVIOS.
- ⁸ Sauer, M., Brown, C., Cavalier, A., Wyatt, C. The Sound-to-Speech/Graphics System for students with severe handicaps. Presentation at RESNA Conference: Kansas City, MO (June, 1991).
- ⁹ Carlson, G. & Bernstein, J. (1988). *A voice input communication aid*. (SRI Project 1252) Bethesda, MD: National Institute of Neuro. & Comm. Disorders & Stroke.
- ¹⁰ Lee, W., Blackstone, S. & Pooch, G. (1987). Dysarthric speech input to expert systems, electronic mail and daily job activities. *Proceedings of AVIOS*. San Jose, CA: AVIOS.
- ¹¹ Smith, D. & Russell, M. (1987). *A study of the performance of automatic speech recognizers with speech impaired children*. Royal Signals and Radar Establishment: No. 4040.
- ¹² Rodman, R.D., Moody, T.S., Price, J.A. (1985). Speech recognizer performance with dysarthric speakers: a comparison of two training procedures. *Official proceeding of Speech Tech '85*, New York, NY, 22-24.
- ¹³ Coleman, C. & Meyers, L. (1991). Computer recognition of the speech of adults

ACN Notes

1. **Having a conference or a workshop?** We will provide you with materials about ACN to hand out. Contact us for details.
2. **ACN can help you teach a class or workshop.** We will put together a special Educational Package/book made up of back issues you choose for your students. 70% discount. Contact ACN.
3. **ACN is copyrighted material.** We ask you to please respect our right to be creative and publish this news report. It is illegal to copy an issue for your colleagues/friends without our permission. If you want additional copies, just call us. We have special multiple copy rates! We'd also like to thank those who already take advantage of our offer and purchase multiple copies.

with cerebral palsy and dysarthria. *Augmentative and Alternative Communication*. 7:1, 34-42.

- ¹⁴ Treviranus, J. Shein, F., Haataja, S. Parnes, P., Milner, M. (1991). Speech recognition to enhance computer access for children and young adults who are functionally nonspeaking. *Proceedings of the 14th Annual Conference*. Resna Press: 308-310.
- ¹⁵ Jutta Treviranus. (December, 1991). Personal communication.
- ¹⁶ Goodenough-Trepagnier, C. & Rosen, M. (1991). Towards a method for computer interface design using speech recognition. *Optimal integration. Proceedings of the 14th Annual Conference*. Resna Press: 328-329.
- ¹⁷ Cheryl Goodenough-Trepagnier. (December, 1991). Personal communication.
- ¹⁸ Tyler Carpenter. (December, 1991). Personal communication.
- ¹⁹ Pooch, G. (1991). Know how to make your speech recognizer work. *Proceedings of the Sixth Annual Conference Technology and Persons with Disabilities*. California State University- Northridge: 771-778.
- ²⁰ Alan Newell. (January, 1991). Personal communication.

RESOURCES

Susan Blockberger, Sunny Hill Hospital For Children, Vancouver, British Columbia Canada (604) 434-1331

Diane Bristow, CSUN, Northridge, CA (818) 885-2869.

Tyler Carpenter, Institute on Applied Technology, Children's Hospital, Boston, MA (617) 735-6998.

Cynthia Cress, Trace Research and Development Center, University of Wisconsin, Madison, WI (608) 262-6966.

Terry Gandell, McGill University, 6275 Northcrest #1017, Montreal, Quebec, Canada H3S 2N3

Cheryl Goodenough-Trepagnier, Department of Rehabilitation Medicine, Tufts University School of Medicine, Boston, MA (617) 956-5036.

Jeff Higginbotham, State University of New York-Buffalo, Buffalo, NY (716) 636-3400.

Sheri Hunnicutt, The Royal Institute of Technology, Stockholm, Sweden. +46 790-7874.

Janice Light, Penn State University, State College, PA (814) 863-2010.

Pat Mirenda, Professor, University of Nebraska, Lincoln, NE (402) 472-4922.

Millet Learning Center staff, Saginaw MI (517) 777-2520.

Caroline Musselwhite, Private Practice, Litchfield Park, AZ (602) 935-4656.

Alan Newell, Microcomputer Centre, University of Dundee +44 382 23181.

Judy Seligman-Wine, Ramat Motza, Jerusalem, Israel +972 2 346078.

Howard Shane, Communications Enhancement Clinic, Children's Hospital, Boston, MA (617) 735-6466.

Sachi Tamura, 467 Fraser St. Victoria, BC, Canada V9A 6H2

Jutta Treviranus on behalf of the Augmentative Communication Services staff, Hugh McMillan Center. ((416) 424-3805.

Tony Vitale, Assistive Technology Center, Digital Equipment Corporation, Northboro, Massachusetts (508) 351-5233.

Michael Williams, Box 10098, Berkeley, CA 94709.