Prediction and Conversational Momentum in an Augmentative Communication System

Norman Alm

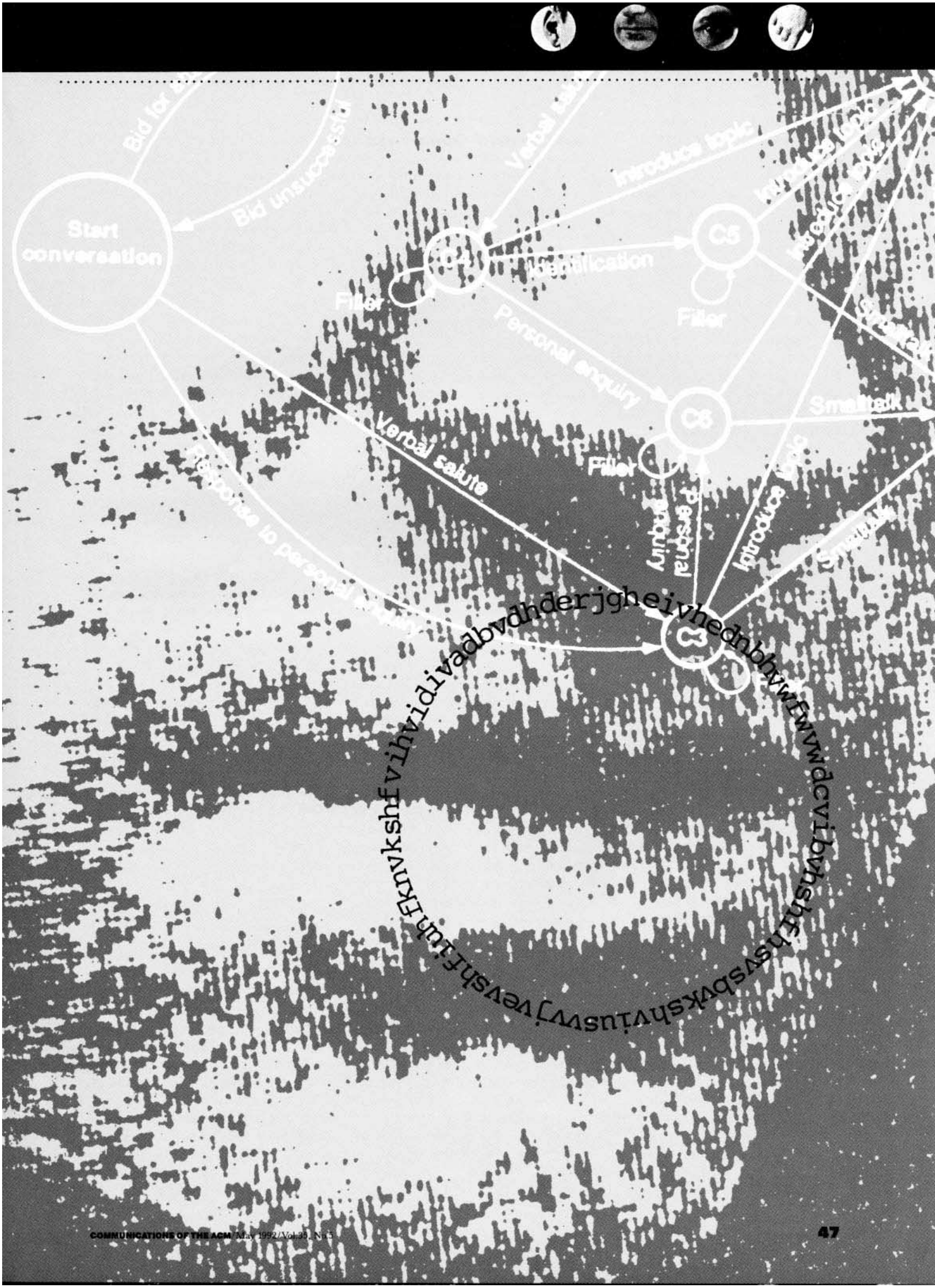
John L. Arnott

Alan F. Newell

Current computer technology offers severely physically impaired nonspeaking people the potential to communicate by using a microcomputer or dedicated microprocessor to control a speech synthesizer. There are several commercially available devices of this type. There remains, however, a serious problem of communication rate, even using the help afforded by existing systems. Given the general physical dysfunction which usually accompanies loss of speaking ability, most nonspeakers have great difficulty in making quick activations of any sort of input device to the computer system. With current systems, users can typically only attain rates of less than 15 words per minute, and many are considerably slower than this. This rate makes it possible, for instance, to write a letter, but it is substantially slower than that needed for taking a full part in conversations that normally proceed at 150–200 words per minute [15].

Communicating with a substantial degree of effort at less than 15 words per minute allows for message passing. Particularly with listeners outside the person's close circle, however, it is extremely difficult to accomplish such vital conversational purposes as projecting one's personality, acquiring a feeling of belonging, creating and changing others' perceptions of the speaker, and influencing and controling others' behavior. These goals are central to human communication, and may be the only purpose of some conversations. The problem for a nonspeaker is that this type of communication is crucially dependent on timely intervention into the conversation, and on the richness and detail of what is said. When a tremendous effort and a great deal of time are required to create even a short message, this kind of social interaction is extremely difficult.

Research in human communication has demonstrated the negative effects of silence [32] and of slow communication rate [38] on the listener's attitude to a speaker. Even with speakers whose physical disability is an obvious cause of slow communication rate, studies have demonstrated serious interactional difficulties. One study has found that below three words per minute,



Start conversation
Bid unsuccessful
Verbal salute
Introduce topic
Identification
C5
Filler
Personal enquiry
C6
Smalltalk
Response to personal enquiry
Personal enquiry
Introduce topic
C3

listeners find such communication to be intolerable [16]. Other research has demonstrated that, with conversational pairs of communication aid users and natural speakers, the natural speakers tend to dominate the conversation [14, 27, 28]. Although there is a wide individual variability in the communicational skills of nonspeakers, research findings give weight to the plentiful anecdotal evidence of misattribution which nonspeakers encounter. They can experience being treated as having low intelligence, spoken to as if they were deaf, or perhaps worse, treated as if they were not even present [13, 29].

This article describes a prototype augmentative communication system designed to create opportunities for a nonspeaking person to take part in the kind of conversational encounters where an almost normal communication rate and fluency are essential for success. The intention of the design is to minimize the effort needed to create and output appropriate utterances.

Applications of Conversation Analysis and Dialogue Design

A prototype communication system has been developed which incorporates principles from conversation analysis and dialogue design, and uses a model of typical conversational patterns to predict the next conversational act the user may want. It provides the user with what can be described as increased conversational momentum, that is, a small effort can initiate relatively long conversational moves. The conversational momentum is produced by this prediction facility, and by an emphasis on performing speech acts. Rather than the user selecting words or specific sentences, the system works by offering the user entire speech acts as basic units for selection. When predicting the next conversational move and when selecting an utterance to accomplish a speech act for the user, the system operates with partial autonomy, but within acceptable guidelines. One guideline is the set of social rules governing conversational interaction. Another is that all the text available is created by the individual user, in his or her own time, for outputting when needed on future occasions.

Research into conversation and human/computer dialogue design normally assumes that participants in the interaction have knowledge about what the other participants are saying.

If the computer system's role in the interaction is to augment the communication of one of the speakers, however, the situation is quite different. Here the system is a tool being used by one of the human participants. The computer system in this circumstance will only have information about one half of the dialogue: the user's utterances thus far. In practice, however, it has been found that this is usually sufficient to provide significant help to the user. Most conversational interaction has a symmetry and this means that the models representing each participant's conversation will have similar structures. Also, such an augmentative system is under the general guidance of the user, and is not required to produce the entire conversation autonomously.

Even with the above caveat, however, the prototype described in this article embodies techniques which are of general interest in the field of dialogue design. A tool such as this for generating conversational output in a semiautonomous way may also be used in exploring a range of human/computer interaction issues.

Predicting Conversational Moves

The prototype communication system attempts to predict a speaker's next utterance. At first consideration, it may seem that predicting conversational acts is not possible. Language is infinitely creative and variable. In fact, Langendoen and Postal argue in their "natural language vastness theorem" that the collection of sentences in any language is so large that it cannot be described by any number, finite or transfinite [24].

Speech does seem to be infinitely variable. It does not follow, however, that it is totally unpredictable. The examples used in traditional linguistics usually are artificially constructed sentences and are normally considered in isolation. This approach may be appropriate where the goal is to generate and pass judgment on individual sentences. Stubbs, however, has commented that "Chomskian linguistics has often over emphasized the creativity of everyday language. In practice, a significant percentage of conversational language is highly routinized into prefabricated utterances" [41]. Fillmore states this more strongly: "an enormous amount of natural language is formulaic, automatic, and rehearsed, rather than propositional, creative, or freely generated" [19].

If a complete conversation is considered, at the most simple abstract level it can be said to have three components, in the following order:

(1) Opening the conversation
(2) Conducting topic discussion, and
(3) Closing the conversation

Opening the Conversation

The process of opening a conversation has the following possible elements, in the order given below [see 25, 35]:

a. Bid for attention
b. Verbal salute
c. Identification
d. Personal inquiry, and
e. Smalltalk

Conducting Topic Discussion

Although topic discussion contains some predictable sequences and routines, it has a much less obvious structure than the opening and closing sections of a conversation. Research in discourse analysis has only just begun to map out the wide and varied area of topic discussion [20, 21, 26]. It may be possible, however, to take Stubbs's and Fillmore's view about the amount of nonoriginal material included in everyday conversation as a starting point and to conclude that what can be predicted about topic discussion is that it is often repetitive.

Closing the Conversation

Just as in opening a conversation we employ a predictable routine; we also tend to follow a set procedure to close the conversation. As Schegloff and Sacks state it, a conversation "does not simply end, but is brought to a close" [37].

The basic elements, in order, of closing a conversation are [see 23, 25, 26]:

a. Transition signals
b. Exchange of phatic remarks, and
c. Exchange of farewells

A Conversation Model

With these features as guides, a model of conversation has been designed which is simple enough to account for a wide range of dialogues, yet contains enough detail to offer the possibility of generating satisfying and effective conversation through predicted sequences.

The formalism used for the model is an augmented transition network (ATN) [47, 30, 34]. The states in the network represent stages the conversation can reach. The transitions between the states occur when a particular event happens, usually the production of appropriate speech acts. The model, although thus far fully realized in only the opening and closing stages of a conversation, does attempt to picture an entire conversation, and not just one portion of it. It also allows for several types of conversation. It accounts for the conversational output of one participant only, but the same model can be used to describe the output of any of the participants in a conversation, with the models interacting by moving together in parallel through the conversational encounter.

A diagram of the model is shown in Figure 1. The states of the network represent stages in the conversation which are reached by sequences of events. Following the ATN formalism, an arc in the network can represent a simple event or a subnetwork of events. Traversing an arc is accomplished either by the occurrence of the given event (normally the production of the relevant speech act), or by jumping to the subnetwork indicated and returning when that subnetwork has been traversed. In this version all states are considered as candidates for 'final' states, which allows maximum flexibility in representing conversations.

From the start state, the conversation can move on by means of a bid for attention, a verbal salute, or a response to another's personal inquiry. If the first act is an attention bid, the conversation only moves on if the bid is successful. The verbal salute may be followed by a personal inquiry, some smalltalk, or the speaker may move straight into a topic introduction. It should be noted that from any state there is the option of using filler remarks, which maintain the speaker's participation in the conversation but do not move it on to a further stage.

Providing Additional Conversational Momentum

In addition to a prediction facility, another way to increase the speaker's conversational momentum is to continually provide the opportunity to output entire speech acts with one keystroke. The identification of a speech act as a unit of discourse is generally regarded as having begun with Austin's exploration of utterances which in themselves constituted an action [4]. A phrase such as 'I hereby pronounce you husband and wife' cannot be interpreted in terms of its truth or falseness. Instead it is one of a class of utterances which, being said, have the power to alter the social environment.

An entire tradition of dialogue analysis has grown from these beginnings, and even though there are unresolved issues within the field, some of its insights could apply to future developments of augmentative communication systems. Thus, as augmentative communication systems incorporate more complex information, contributions such as Searle's work on felicity conditions and indirect speech acts [39], Grices's work on implicature [17, 18], Sperber and Wilson on relevance [40], and Cohen, Perrault and Allen on plan recognition [1, 11, 33], may be usefully applied to them.

The prototype described here makes use of the initial observation that most things people say to each other are intended to affect the relationship between speaker and listener, rather than being just 'messages' passed between people. It is based on the importance of speech as a vehicle simply of participating in a satisfying way in social encounters. It is also recognized that there tends to be a social purpose underlying even the transfer of objective information. These include being helpful, appearing intelligent, or asserting one's own views and prejudices on the subject. The implication for an augmentative communication system is that it would be very economical of effort, and more likely to help in making real conversation, if one of the system's basic units was the speech act, and if appropriate speech acts could be made easily available to the user.

The vast possibilities of natural language mean that, even at the speech act level, there are a great many possibilities, each with many different modes of expression. The set of speech acts concerned with greeting and departing rituals tend to be fairly predictable with the same individual. On the other hand, the speech acts appropriate to topic discussion are potentially a very large class, and do not always appear in the same predictable sequences. The ATN model we are using handles this uncertainty at present by employing a subnetwork

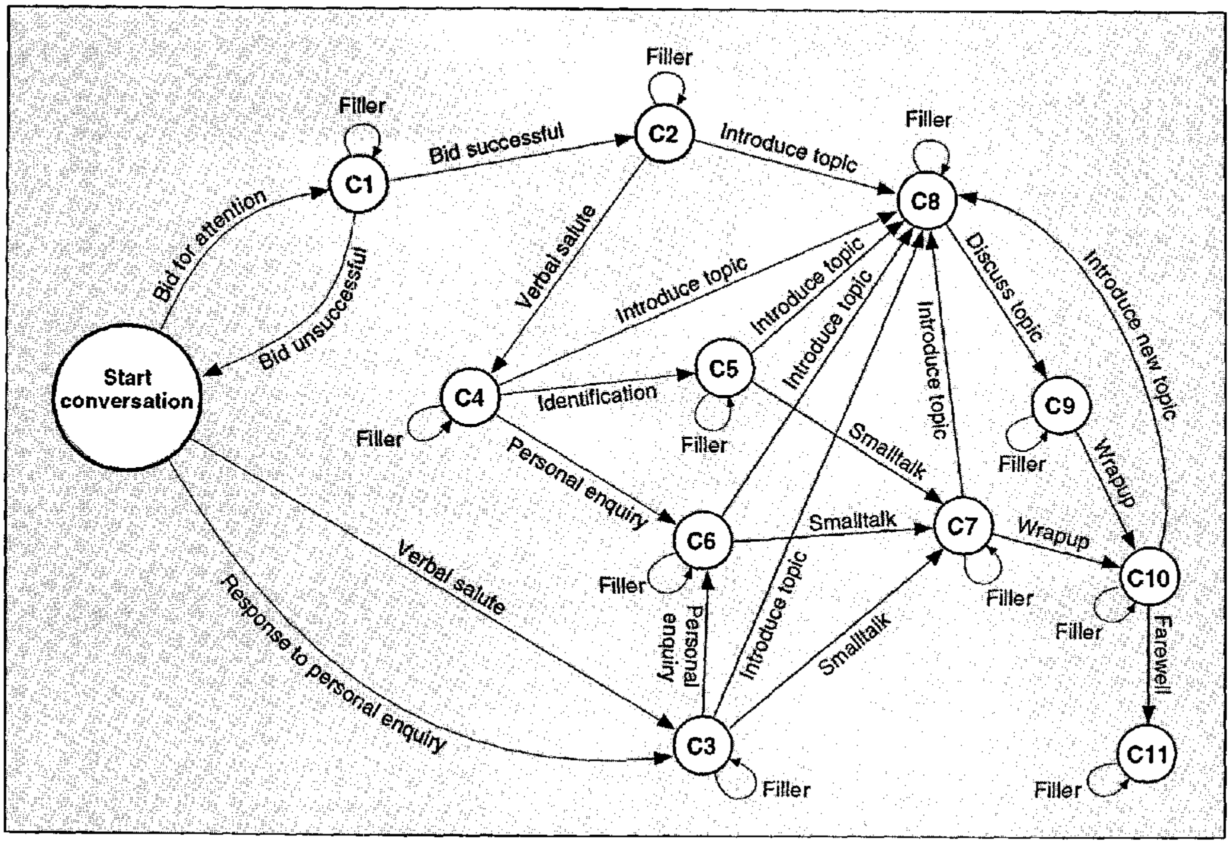


Figure 1. The role of one participant in a conversation modeled with an augmented transition network

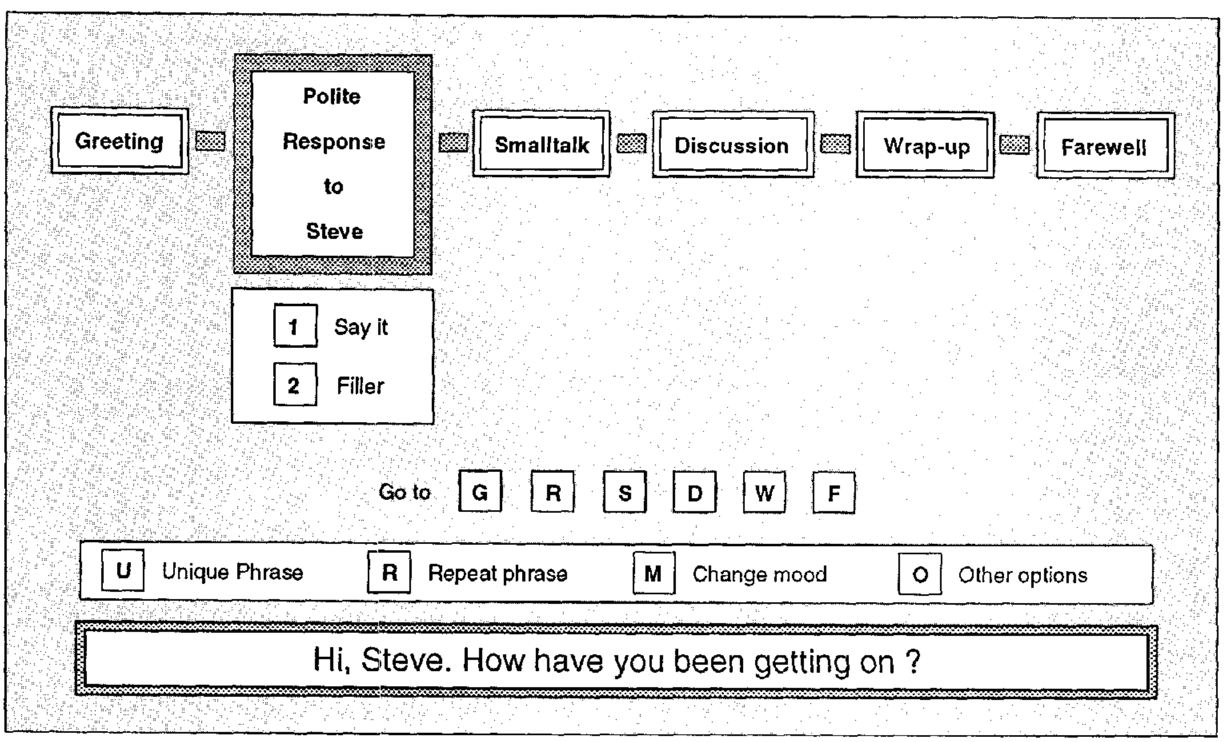


Figure 2. CHAT control screen for opening and closing a conversation

for topic discussion which has a very general structure, and no predictive abilities.

A number of approaches are being taken to introduce more structure into the topic discussion stage of this model of a conversation. It may be possible to predict a set of speech acts that will be particularly appropriate for the given situation. A close study has been undertaken of related research done on natural language interfaces to databases [44, 45], linguistic work on the sequentiality of speech acts in naturally occurring conversation [22], and the use of scripts for encapsulating stereotypical social patterns [36].

Also being explored is the possibility of helping a user to include reusable conversational texts easily in a conversation. For example, a significant portion of conversational output is devoted to narratives. These are often repeated more or less verbatim with different people, and include material such as recent family news, favorite jokes, the procedure for locking the wheelchair, how the vacation went, and so on.

An important and reasonably small class of speech acts used in topic discussion is the set of utterances used to give feedback to another speaker. It is an oversimplification to view conversation as a sequence of turn-taking of speaking and listening, as conversation tends to falter without active participation by the listener. A conversation is thus an event that is created simultaneously by all the participants. The speaker needs to monitor continuously whether and how the intended meaning is being perceived. The listener needs to convey reactions to what the speaker is saying, or signal a wish to take a speaking turn, while not unduly interrupting the other speaker's flow [7].

Prototype Discourse Management System

A prototype communication system, called Conversation Helped by Automatic Talk (CHAT), has been developed, which embodies much of the conversation model described, and incorporates a number of the principles we have just discussed. The intention of the design is to have the system automate as much of the conversational process as is feasible, and in this way reduce the keystrokes necessary to operate it, substantially increasing the rate of conversational participation that is possible by a nonspeaking person.

The CHAT system was written in Pascal. It produces satisfactory real-time performance with no processing delays. The hardware platform is a laptop PC, with a text-to-speech synthesizer. This configuration is widely available at a reasonable cost.

CHAT potentially gives a nonspeaking person the same type of adjustable control as a walking stick gives someone having difficulty with walking. The stick can bear as much, or as little, of the user's weight as necessary, depending on the user's situation. The CHAT user can choose to have the system say something just by a command that amounts to instructing it to "Say something sensible here." This can be accomplished with a single keystroke (or equivalent action on a severely impaired user's special switch). More precise control over the output can be exercised, with a concomitant increase in selection effort, up to the final option of creating a unique message.

The CHAT prototype's predictive ability is thus far employed only in the opening and closing stages of a conversation, and in the set of feedback remarks that are presented in the topic discussion phase. The opening and closing stages of a conversation are largely ritualized encounters and thus amenable to implementation with the version of the ATN model we have described. The set of feedback remarks necessary to cover a large number of discourse situations is not large. As stated previously, further research is being undertaken to discover methods of applying predictive techniques to other parts of the topic discussion phase of a conversation.

The CHAT prototype operates by consulting its model of a conversation, and offering the user predictions about the next conversational move. These are specified in terms of speech acts, not specific utterances. Since CHAT operates at the level of a speech act, and not a specific utterance, it is able to automatically provide a variation in output, simulating what unimpeded speakers do to avoid clumsy repetition. The user can opt for CHAT's predicted speech act, or direct it to output another type of speech act. Where CHAT is unable to help, the user always has the option to create unique text (at, of course, a much slower rate).

Another feature that helps to encourage natural conversation is mood specification: the user can specify a mood, then all subsequent utterances from CHAT reflect that mood. For instance, the user may begin in a polite mood, but be angered by something the other person says to them. The angry mood can then be easily set, and the phrases that are output will be in an angry form, in addition to being appropriate for the current stage and conversational move. An informal mood can be used for close friends.

A stored list of the names of known people is held, which allows the automatic insertion of the conversation partner's name at appropriate points in the utterances. This list can include information on the mood setting wanted for particular individuals, so that, for instance, a close friend would always be greeted in an informal manner. Other mood choices might be humorous, sad, sarcastic, playful, or whatever suits the user. The structure of the CHAT system means that all that is required is for a list of phrases in each of the mood choices to be stored in the CHAT system for all of the conversational possibilities.

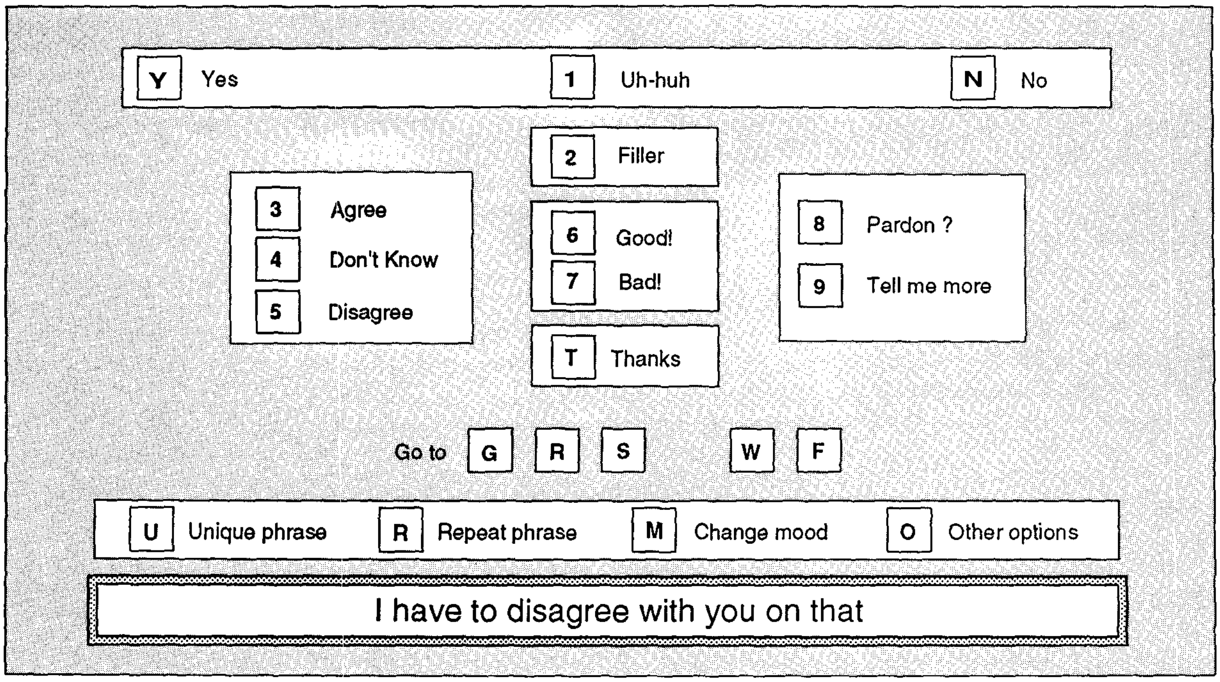


Figure 3. CHAT discussion stage control screen

The user can originate all of the stored phrases, entering them in his or her own time, when input rate is not important. The stored text can be changed or updated simply. The intention is for the phrases to have the personal stamp of the user, and to be modified over time, according to changing needs and preferences. This will help to counteract any tendency towards impersonality when a speech act is automatically selected during a conversation. Through this procedure, all the possible choices will reflect the personality of the user, even though the selection of a particular utterance is made by the computer.

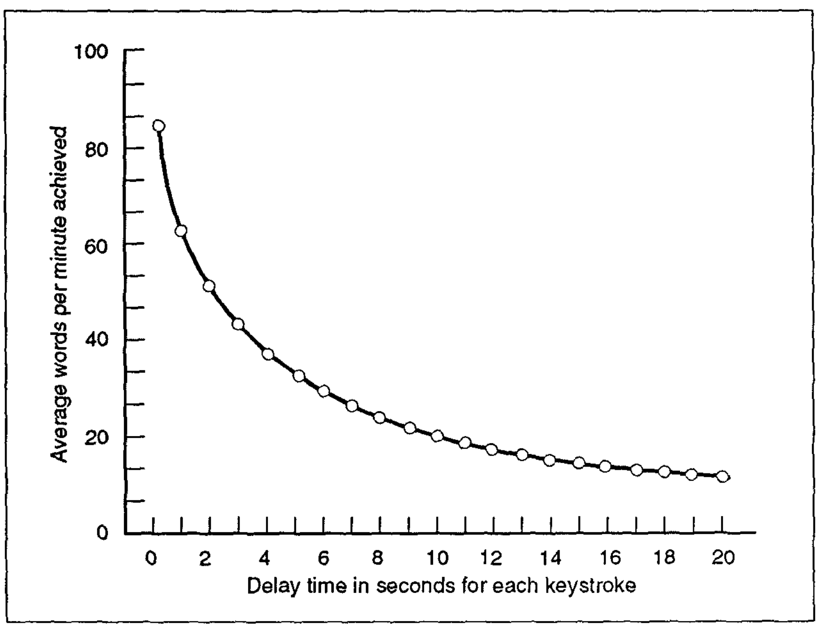


Figure 4. CHAT words per minute rates when a range of delays in keying time is introduced

Opening and Closing Stages of Conversation

The control screen for helping the user navigate through the opening and closing stages of a conversation is shown in Figure 2. The user is speaking to a conversational partner called Steve, and has just given Steve a Polite Greeting, by pressing Key 1 (Say it) with the prediction window over Greeting. The system has selected the Greeting "Hi, Steve. How have you been getting on?", and output it. The prediction window has now moved on to predict that a response to a greeting from the other person will be wanted next.

The system offers a total of twelve choices for the next action. Thus, for example, by pressing keys G, R, S, W, F the user can direct the system to speak a greeting, give a response to the greeting, make a smalltalk remark, and so on, rather than use the speech act predicted. Also the last phrase can be repeated, or the mood of subsequent utterances can be altered.

Filler Remarks

The CHAT system allows the user to at any point output a filler remark. Filler remarks are innocuous, multi-purpose utterances whose primary function is to convey the speaker's bid to maintain a place in the interaction, without conveying any substantive meaning.

Filler remarks are stored and output in the same way as any of the other phrases. The stage of the conversation will determine the appropriate list of possible fillers, and a constrained random choice is made from this list. Fillers at the greetings stage will include phrases such as "Good to see you", as well as more generic utterances, such as "So . . .". Similar rules will apply to other stages. Different moods also require different fillers. The angry mood setting will produce terse, neutral filler remarks, for instance.

Feedback Remarks

Continuous feedback from the listener to the speaker is important in creating the rapport necessary for conversation. This feedback can, as with fillers, merely denote continued attention. It can also convey more information, such as agreement, puzzlement, amusement, shock, or other reactions.

A range of feedback remarks have been included in CHAT, which are available, in the Discussion stage of the conversation, for the equivalent of one keystroke. Having only one keystroke per speech act is particularly important when giving feedback to another speaker, since timing is important for the remark to have its effect. The feedback remarks were chosen to cover as wide a range of situations as possible. The control screen for the Discussion stage of conversation is shown in Figure 3.

It must always be possible for a user to create unique text, however well a predictive system operates, although this produces a significant time penalty. There are input acceleration techniques that can assist with this. When a user wishes to create a unique utterance in the current CHAT prototype, he/she presses key 'U' (from either control screen) and assistance is offered by the Predictive Adaptive Lexicon (PAL) system, that was developed at the University of Dundee. PAL is a word prediction system based on frequency and recency of word usage. It typically saves 50% in keystrokes needed to create unique text [42].

Field Testing

The CHAT system described earlier was implemented on an MS/DOS machine and initially pilot-tested with nonimpaired subjects [2]. Four physically impaired nonspeaking people then used it on a trial basis [3, 8, 9].

These tests were designed to determine the feasibility of the CHAT concept, and focused on the types of conversation in which the speaker can participate, employing just openings, smalltalk, feedback remarks, and closings. The results were positive in terms of the three factors which constitute the principal weak points of existing systems. These are

(1) speed of communication,
(2) naturalness of the conversation produced, and
(3) the ability to easily express mood and individual personality.

Speed

The CHAT system was found to provide a much faster method of generating this type of conversational material than existing devices. Naturally spoken conversation proceeds at 150–200 words per minute, whereas Foulds reported typical rates achieved by communication device users of below 15 words per minute [15]. Newer devices have since produced a number of improvements, but this estimation of achievable rates still applies [2]. A simulation of the effect of a range of physical disability in controlling CHAT can be made by using data from the pilot study with nonimpaired subjects and adding a nominal delay time for each keystroke. The word rate achievable with CHAT was calculated using this method at 12 to 85 words per minute, depending on the degree of simulated disability. The results are shown graphically in Figure 4. The results from tests of CHAT with four physically impaired nonspeakers, who achieved respective average rates of 19, 28, 42 and 54 words per minute, confirmed these calculations.

Simulation of Natural Conversation

Because of their slowness in controlling an input device, severely physically impaired people with currently available systems are unable to include in their communications the sort of noninformational content that is common in unimpeded speech. All of their effort is needed to produce a short message that conveys the information they want to impart. CHAT allows the user to produce socially bonding speech acts which are crucially dependent on quick and timely production for their effect.

The user can also create and modify phrases in CHAT, when conversation is not occurring, so CHAT can easily contain slang, humorous turns of phrase, and casual phrases of the user's own making. The inclusion of the speaking partner's name, and the ready availability of filler remarks, also help to make the dialogue more natural and interactive.

The CHAT system helps a user to make conversation which sounds more natural, and therefore more acceptable to people who may not know the user well, and who may not be aware of the severe limitations on communication that the disability imposes. Even with conversation partners who are quite familiar with the user, the ability to participate more fully in normal, verbal interaction was found to be of great value.

The transcripts given next show the naturalness of the dialogue achieved by the CHAT prototype in the pilot tests performed with non-impaired users.

Extracts from test runs of CHAT:

A: Hello, Bill. How are you today?
B: Fine, and yourself?
A: Just fine thanks, Bill. Not bad at all. How have you been getting on?
B: Quite well, and yourself?
A: Och aye, well. What's the latest news? Anything interesting?
B: Oh, the usual dull routine, I'm afraid. Nothing much to tell you.
A: Uh-huh.
Time taken: 55 seconds
Mood: Polite

A: Hi, Ann. How are you?
B: I'm fine thank you. How are you?
A: Not bad thanks, Ann. We're moving along as usual here.
B: That's good.
A: How have you been getting on?
B: Very well, thank you.
Time taken: 25 seconds
Mood: Polite

A: What happened, I mean, I noticed you weren't there.
B: I don't really know.
A: Something to do with transport maybe?
B: Yeah. Right.
A: Oh. Ah well. At least you didn't miss too much. I mean there wasn't anything exciting happened anyhow.
B: Oh good.
A: [Laughs]. Ah well, let's hope nothing too bad happens the next time—you get along the next time OK.
B: Yes. That's great. Thanks. So. That's how it is, I suppose.
A: OK then. [Laughs].
Time taken: 39 seconds
Mood: Angry

In the first two extracts speaker A is using CHAT and in the third it is speaker B. Both subjects in each case were able-bodied. These dialogues take place in something approaching a natural turn-taking rhythm. When CHAT was tested with nonspeakers, both users and conversation partners reported in a questionnaire that the dialogues seemed more natural than ones held without using CHAT.

Expressing Emotion and Personality

The facility to change the mood of the phrases is an important feature of CHAT. This enables the user to express feelings and personality in a way that is extremely difficult for current communication system users. Since all stored phrases will be in the user's own style of expression, it ensures that whatever is said will contribute to projecting his or her personality. In the questionnaires following tests with nonspeakers using CHAT, the conversation partners said they felt the users' personalities and moods were conveyed more strongly when

Input acceleration techniques for severely physically disabled nonspeakers using a communication system

A number of methods have been devised to accelerate the input process for a severely impaired nonspeaker using a communication system. A certain amount of success has been achieved. There remains, however, a large gap between the typical rates achievable by a nonspeaker using existing systems and unimpaired speakers. Typically, rates for nonspeakers using existing systems are 2–10 words per minute, whereas unimpeded speech proceeds at 150–200 words per minute. It is the overall quality of communication that is important, not just the rate. For satisfactory interaction, particularly with uninitiated conversation partners, a very slow rate is a severe impediment.

There are three modes of input used with communication systems relative to the user's physical control difficulties. These are direct selection, coding, and scanning. With a direct selection aid, the user can point to his or her choice from a selection, or can use a keyboard or similar device. The coding and scanning modes are usually employed along with specialized individual switches. Since lack of speech is usually accompanied by general physical impairments, nonspeakers often require special switches to act as input devices to control a communication system. A wide range of switches and sensors provides users with a way to operate the system by means of whatever body movement they can best control. Coding systems allow for the use of a small range of movements to control the communication aid. Morse code, for instance, has been employed where a user can control both hands, but can make only a small amount of movement with them. One hand can then activate a 'dot' switch and one a 'dash' switch. Where movement is very severely restricted because of the user's physical impairment, a scanning system may be used. Here, selection of items from an array of elements is accomplished by any action that can effect a switch closure or sensor activation. It is then necessary to have a system scan through the choices possible in order to close in on the user's choice. A number of scanning strategies may be employed to make this process more efficient; for instance, row-column scanning, and quarternary or binary scanning.

Even when the switch and input mode has been optimized for a particular user, communication is still extremely slow. A variety of methods have been developed to accelerate the input process. Acceleration methods thus far developed fall into two categories: achieving more output for a minimal input, and applying linguistic and other knowledge in order to predict what the user might want to say/write next.

using CHAT than when using their usual mode of communication.

The strongest negative comment in the questionnaire from nonspeaking CHAT users was that it does not help them in creating topic-based discussion. At present, CHAT requires that the user employ whatever conventional methods they already use to convey unique information. As reported, current research is addressing this problem.

Conclusion

The CHAT prototype demonstrates that in some of the most common conversational situations, real time discourse management using a computer-based communication system is possible, requiring only minimal control actions from a user. The application of this technique to communication systems for nonspeaking, severely physically impaired people allows them to participate more successfully in some conversational interactions at near to normal rates. Future research will attempt to identify discourse patterns in topic discussion, which will allow the system to increase the help it can offer in this aspect of conversation.

The tests with CHAT have shown that effective, natural-sounding conversation is possible between two people, one of whom is 'speaking' via a speech synthesizer and using a computer-based discourse management system with an inbuilt conversational model to provide predictions and increase the conversational momentum. The results from this study will lead toward the development of a general purpose speech prosthesis for physically impaired nonspeakers. In addition, however, such results and techniques are potentially applicable to those situations where more natural dialogue structures are seen to be desirable within human/machine interaction. ◼

Minimal Input/Maximum Output

Storing entire phrases which the user could access with a small number of activations was an early proposal for rate acceleration. Certainly, with very common phrases such as "How are you?" this can work, but the problem is that phrases are very context-dependent, and a very large number are needed to allow for many situations. Remembering and accessing a very large number has not proved practicable. Using prestored material emphasizes fluency over specificity, and users seem thus far to prefer to struggle on with being specific in their utterances despite the rate problem.

Another method is to use abbreviations for common words and phrases, and have the system expand the abbreviated input [43]. There is a ceiling on the improvement possible in using abbreviations for stored words, however, related to the average word size in the given language. For large numbers of such abbreviations, there is a problem of cognitive load for the user, in remembering both what items are stored and their individual access codes.

A word-coding system based on icons that can have multiple meanings is currently available [5]. For common words, given a small number of icon sequences, this represents a savings in keystrokes needed to produce the word. The icons have the additional advantage of meaning, which helps the user to remember sequences, and they may also be used as a communication system for nonspeakers who are also nonreaders.

Prediction Using Linguistic Information and User Monitoring

Human language incorporates a great deal of redundancy, and this is exploited in systems that predict the word(s) the user is attempting to create with a communication system. Word completion systems depend on information about word frequencies and the user's previous selections [42]. Word pair prediction depends on logging the user's output, and offering likely candidates for the next word or words [6, 46].

The current situation is that some help is available for severely impaired nonspeakers, using the methods just described, but existing communication systems for such users are still extremely slow in output, and do not enable their users to approach, even remotely, normal conversation word rates.

One current research effort in minimal input methods is directed toward producing a system which could take as input a few key words, and expand this into the sentence the user wanted to type [12]. Attempts are also being made to improve prediction methods through the incorporation of more linguistic information [31, 48]. A third approach is to model some of the semiautomatic and reusable features of everyday conversation in a communication system. This is the method described in the accompanying article.

—Norman Alm, John L. Arnott, and Alan F. Newell

References

1. Allen, J. Recognizing intentions from natural language utterances. In Brady, M., and Berwick, R.C., eds., *Computational Models of Discourse.* MIT Press, Cambridge, Mass., 1983.
2. Alm, N. Towards a conversation aid for severely physically disabled non-speaking people. Ph.D. Thesis, Dept. of Mathematics and Computer Science, University of Dundee, Dundee, Scotland, 1988.
3. Alm, N., Brophy, B., Arnott, J.L., Newell, A.F. Preliminary evaluation of a conversation aid based on speech acts. Abstract in *Augmentative and Alternative Commun. 4,* 3 (Sept. 1988), 153.
4. Austin, J.L. *How to Do Things with Words.* Clarendon Press, Oxford, 1966.
5. Baker, B. Minspeak. *Byte.* Sept. 1982. 186–202.
6. Bentrup, J. Exploiting word frequencies and their sequential dependencies. In *Proceedings of the 10th Annual Conference on Rehabilita-*

tion Technology (San Jose, Calif., June 19–23). RESNA–Association for the Advancement of Rehabilitation Technology, Washington, D.C., 1987, 121–123.

7. Blau, A. Communication in the back-channel: Social structural analyses of nonspeech/speech conversations. Ph.D. Dissertation, City University of New York, 1986.
8. Brophy-Arnott, M.B., Alm, N., Newell, A.F., Arnott, J.L. The effect of a predictive communication aid employing speech acts on the conversation of non-vocal speakers. Paper submitted for publication, 1991.
9. Brophy, M.B., Colvin, J.B., Fuller, D., Swiffin, A.L. Communication aid for the neurologically impaired child and young adult in the school and community. Final Report for the Scottish Home and Health Department, Grant no. K/RED/4/C56, University of Dundee, Dundee, Scotland, 1989.
10. Brown, A. Inputs and how to activate a communication aid. In Enderby, P., ed., *Assistive Communication Aids for the Speech Impaired.* Churchill Livingstone, New York, 1987.
11. Cohen, P.R., and Perrault, C.R. Elements of a plan-based theory of speech acts. *Cognitive Sci.* 3, 177–212.
12. Demasco, P., McCoy, K., Gong, Y., Pennington, C., and Rowe, C. Towards more intelligent AAC interfaces: The use of natural language processing. In *Proceedings of the 13th Annual Conference on Rehabilitation Technology,* (Washington, D.C., June 15–20). RESNA Press, Wash., D.C., 1989, 141–142. (Also see article in this issue on this research.)
13. Eulenberg, J. (ed.) *Conversations with Non-speaking People.* Canadian Rehabilitation Council for the Disabled, Toronto, Canada, 1984.
14. Farrier, L., Yorkston, K., Marriner, N., and Beukelman, D. Conversational control in nonimpaired speakers using an augmentative communication system. *Augmentative and Alternative Commun. 1,* 2 (May 1985), 65–73.
15. Foulds, R. Communication rates for non-speech expression as a function of manual tasks and linguistic constraints. In *Proceedings of the International Conference on Rehabilitation Engineering,* (Toronto, Canada, June 16–20). Rehabilitation Engineers Society of North America, Toronto, 1980, 83–87.
16. Goodenough-Trepagnier, C., Galdieri, B., Rosen, M.J., and Baker, E. Slow message production rate and receivers' impatience. In *Proceedings of the 2nd International Conference on Rehabilitation Engineering* (Ottawa, Canada, June 17–22). Rehabilitation Engineers Society of North America, Bethesda, Maryland, 1984, 347–348.
17. Grice, H. Logic and conversation. (From William James Lectures, Harvard University, 1967). In Cole, P. and Morgan, J., eds., *Syntax and Semantics, v. 3, Speech Acts.* Academic Press, New York, 1975.
18. Grice, H. Further notes on logic and conversation. In Cole, P. *Semantics, v. 7, Pragmatics.* Academic Press, New York, 1978.
19. Gumperz, J. *Discourse Strategies.* Cambridge University Press, London, 1982.
20. von Hahn, W. Pragmatic considerations in man-machine discourse. In *Proceedings of the Eleventh International Conference on Computational Linguistics* (COLING 86). (Bonn, West Germany, Aug. 25–29). Association for Computational Linguistics, Bonn, 1986, 520–526.
21. Hopper, R. Interpretation as coherence production. In Craig, R. and Tracy, K., eds., *Conversational Coherence—Form, Structure, and Strategy.* Sage Publications, Beverly Hills and London, 1983.
22. Jose, P. Sequentiality of speech acts in conversational structure. *J. Psycholinguistic Res. 17,* 1 (1988), 65–88.
23. Knapp, M., Hart, R., Friedrich, G., and Shulman, G. The rhetoric of goodbye: Verbal and non-verbal correlates of human leave-taking. *Speech Monographs, 40* (Aug. 1973), 182–198.
24. Langendoen, D.T. and Postal, P. *The Vastness of Natural Language.* Basil Blackwell, Oxford, 1984.
25. Laver, J. Communicative functions of phatic communion. In Laver, J., ed., *Semiotic Aspects of Spoken Communication.* Edward Arnold, London, 1974.
26. Levinson, S. *Pragmatics.* Cambridge University Press, London, 1983.
27. Light, J. Communicative interaction between young nonspeaking physically disabled children and their primary caregivers: Part I—Discourse patterns. *Augmentative and Alternative Commun. 1,* 2 (May 1985), 74–83.
28. Light, J. Interaction involving individuals using augmentative and alternative communication systems: State of the art and future directions. *Augmentative and Alternative Commun. 4,* 2 (June 1988), 66–82.
29. McGregor, A. I'm in prison. *Commun. Outlook. 12,* 3 (Mar. 1991), 21.
30. Metzing, D. ATN's used as a procedural dialog model. In *Proceedings of the International Conference on Computational Lingustics* (COLING 80). Association for Computational Linguistics, Tokyo, 1980, 487–491.
31. Morris, C., Newell, A.F., Booth, L., and Arnott, J.L. Syntax PAL—A system to improve the syntax of those with language dysfunction. In *Proceedings of the 14th Annual Conference,* (RESNA '91) Washington, D.C., June 21–26. RESNA Press, Washington, D.C., 1991, 105–106.
32. Newman, H. The sounds of silence in communicative encounters. *Commun. Quart. 30,* 2 (Spring 1982), 142–149.
33. Perrault, C.R., and Allen, J. A plan-based analysis of indirect speech acts. *Amer. J. Comput. Ling. 6,* 3–4 (Jul.–Dec. 1980) 167–182.
34. Reichman, R. *Getting Computers to Talk Like You and Me—Discourse Context, Focus, and Semantics (An ATN Model).* MIT Press, Cambridge, Mass., 1985.
35. Rosner, M. Three strategic goals employed in conversational openings. Working Paper No. 46. Fondazione Dalle Molle, Institut pour Les Etudes Semantiques et Cognitives, Universite de Geneve, Geneva, 1981.
36. Schank, R. and Abelson, R. *Scripts, Plans, Goals, and Understanding.* Lawrence Erlbaum, New Jersey, 1977.
37. Schegloff, E. and Sacks, H. Opening up closings. *Semiotica, 8* (1973), 289–327.
38. Scherer, K. Personality markers in speech, in Scherer, K., ed., *Social Markers in Speech.* Cambridge University Press, London, 1979.
39. Searle, J. *Speech Acts: An Essay on the Philosophy of Language.* Cambridge University Press, London, 1969.
40. Sperber, D., and Wilson, D. *Relevance.* Oxford University Press, Oxford, 1986.
41. Stubbs, M. *Discourse Analysis—The Sociolinguistic Analysis of Natural Language.* Basil Blackwell, Oxford, 1983.
42. Swiffin, A., Arnott, J.L., Pickering, J.A., and Newell, A.F. Adaptive and

predictive techniques in a communication prosthesis. *Augmentative and Alternative Commun. 3*, 4 (Dec. 1987), 181–191.

43. Vanderheiden, G., and Kelso, D. Comparative analysis of fixed-vocabulary communication acceleration techniques. *Augmentative and Alternative Commun. 3*, 4 (Dec. 1987), 196–206.

44. Wachtel, T. Discourse structure. LOKI Report NLI-1.1. University of Hamburg Research Unit for Information Science and Artificial Intelligence, Hamburg, West Germany, 1985.

45. Wachtel, T. Pragmatic sensitivity in NL interfaces and the structure of conversation. In *Proceedings of the Eleventh International Conference on Computational Linguistics* (COLING 86). (Bonn, West Germany, August 25–29). Association for Computational Linguistics, Bonn, 1986, 35–40.

46. Woltosz, W. Stephen Hawking's communication system. *Commun. Out. 10*, 1 (Summer 1988), 8–11.

47. Woods, W. Transition network grammars for natural language analysis. *Commun. ACM 13*, 10 (Oct. 1970), 591–606.

48. Yang, G., McCoy, K., Demasco, P. Word prediction using a systemic tree adjoining grammar. In *Proceedings of the 13th Annual Conference*, RESNA '90 (Washington, D.C., June 15–20). RESNA Press, Washington, D.C., 1991, 185–186.

CR categories and subject descriptors: C.3 [**Special-Purpose and Application-Based Systems**]: Microprocessor/microcomputer applications; H.1.2 [**Information Systems**]: User/Machine Systems—*human factors; human information processing;* H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*retrieval models; search process; selection process;* I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems—*medicine and science;* J.3 [**Computer Applications**]: Life and medical sciences—*health*

General terms: Human Factors

Additional Key Words and Phrases: Augmented transition networks, dialogue design, discourse analysis, user interfaces

About the Authors:

NORMAN ALM is the Digital Equipment Corporation Research Fellow at the Microcomputer Centre, at Dundee University in Scotland. His primary research interest is in developing improved communication systems for severely physically impaired nonspeaking people. As well as being a computer scientist, he has worked for a number of years in education and social work. **Author's Present Address:** MicroCentre, Department of Mathematics and Computer Science, University of Dundee, Dundee DD1 4HN, Scotland, U.K. email: nalm@uk.ac.dund.mic (JANET)

JOHN L. ARNOTT is senior lecturer in Computer Science at the University of Dundee in Scotland, and codirector of research on Human-Machine Interaction and Computer Systems in Rehabilitation and Therapy at Dundee University Microcomputer Centre. He has a background of many years research on adaptive interfaces and communication systems for speech, hearing, motor, and language impaired people. **Author's Present Address:** MicroCentre, Department of Mathematics and Computer Science, University of Dundee, Dundee DD1 4HN, Scotland, U.K. email: jarnott @uk.ac.dund.mic (JANET)

ALAN F. NEWELL is NCR Professor of Electronics and Microcomputer Systems and Director of the Microcomputer Centre at Dundee University. His research interests are human/computer interaction, including a special interest in disability, and the use of computers in rehabilitation and therapy. **Author's Present Address:** MicroCentre, Department of Mathematics and Computer Science, University of Dundee, Dundee DD1 4HN, Scotland, U.K. email: afn@uk.ac.dund.mic. (JANET)