A communication system based on scripts, plans, and goals for enabling non-speaking people to conduct telephone conversations. N.Alm, A.Morrison and J.L.Arnott. In: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC 1995),* Vancouver, Canada, 22nd – 25th October 1995, Vol.3, pp.2408-2412.

A Communication System Based on Scripts, Plans, and Goals for Enabling Non-speaking People to Conduct Telephone Conversations

Norman ALM, Arthur MORRISON and John L. ARNOTT

Dept. of Mathematics and Computer Science, University of Dundee, Dundee, Scotland, UK.

The research described here was published in:

Proceedings of the 1995 IEEE International Conference on Systems, Man and Cybernetics, Vancouver, Canada, 22nd – 25th October 1995, Volume 3, pp. 2408-2412. ISBN: 0-7803-2559-1.
INSPEC Accession No.: 5132714. Pub: IEEE, New York, USA.

The IEEE (Institute of Electrical and Electronics Engineers) is at: http://www.ieee.org/

DOI: 10.1109/ICSMC.1995.538142
URL: The DOI bookmark for the article is: http://dx.doi.org/10.1109/ICSMC.1995.538142

ABSTRACT

Current communication systems for severely disabled non-speaking people still fall short of allowing them easily to effect many practical tasks, such as getting something accomplished over the telephone. Given that the accomplishment of many such tasks could be said to be 'script' based, the application of a goals-plans-scripts hierarchy to augmented communication system design might produce advantages. A prototype system tested out this idea with promising results. A further system is under development.

INTRODUCTION

Severely physically disabled non-speaking people face formidable challenges in communicating with others. Even with existing speech output technology, the speaking rate of augmented speakers remains very low, commonly as low as 2-10 words per minute. A number of research projects have been investigating ways of improving this situation, including abbreviated input, predictive methods, and conversational modelling.

Research projects involving prediction and conversation modelling, a number of which have been carried out at Dundee University, have been based on the finding from conversation analysis that conversation does often have a very predictable structure. Previous work has resulted in a prototype communication system called CHAT which assists the user by modelling the opening and closing phases of a conversation, and, for the central portion of the conversation, provides the user with readily available feedback remarks, with which to react to the other speaker [1].

Following the development and testing of this prototype, a number of different approaches have been explored to providing predictive help in the topic discussion phase of a conversation.

A system called TalksBack was developed which predicts sentences from a large prestored collection, with the prediction being based on a model of the user's interests, the interests and social profile of frequent conversation partners, and the topic currently under discussion [2].

Another system, PROSE, was developed to facilitate the telling of stories. Far from being confined to childhood settings, the telling of stories is an integral and essential part of all human communication. A great deal of time in conversation is spent in relating narratives to each other. There is evidence that the telling of stories is the primary way in which we build up a projection of our personality with others, and maintain it [9].

Features from CHAT, TalksBack and PROSE have now been integrated into a single commercially available communication system [4]. Early indications from customers are that these new facilities are much appreciated, and are making a difference in the users' communication abilities.

A new application which is being pursued with the sentence prediction and story-telling systems is to investigate the possibility that they may be useful for people with cognitive as well as communicative problems, such as occurs with aphasia following a stroke. Here, the system would provide the user with a prompt as to what of their personal conversational material they might like to use next, and thus act as a kind of 'cognitive prosthesis'. The systems have thus far been tried out with some people who have aphasia with promising results. A fuller investigation is under way [3].

The work also continues on discovering improved ways to model topic discussion in a communication system. The hope is that useful elements which are developed, such as those described above, can eventually be combined into a multi-level system which can help the user cope with a wide variety of communicational situations [7].

In examining what portions of topic discussion might be predictable, the finding was that a great deal of our conversation is reusable, and the systems described thus far exploit this by providing the user with a large store of personal conversational material, and a way to access the store easily and quickly.

THE APPLICABILITY OF SCRIPTS

Another predictable aspect of conversation is the fact that many dialogues take place in order to accomplish a particular purpose. This applies to commercial transactions, such as dialogues in shops, and also to discussions with professionals such as doctors. In fact any situation in which the speaker is trying to accomplish a particular concrete goal is likely to be predictable, because these situations are rarely unique. They repeat themselves regularly. Just as conversational elements are reusable, so the sets of utterances we use to accomplish a particular purpose are likely to be reusable. They could be said to form a kind of 'script'.

Schank and Abelson developed a theory of scripts based on the observation that people do not enter into every situation completely unprepared for it [8]. As has been noted above, many, if not most situations we encounter call for a certain amount of stereotypical communication. Some situations consist of nothing but this form of communication. A script-based communication system for non-speakers could provide a useful form of prediction for them in situations where they are wanting to accomplish some action. Also the hierarchy which Schank and Abelson developed of

GOALS ⟶ PLANS ⟶ SCRIPTS

provides a method of coping with breakdown, or unforeseen events resulting in the premature exit from a script.

A PROTOTYPE SYSTEM

In order to test the feasibility of using goals, plans, and scripts in an augmented communication situation, a prototype was developed based on this structure. The conversation task selected for the tests was that of the user telephoning to get something done. This type of conversation was chosen because it has a clear goal, with a number of possible plans and scripts to achieve it. For non-speaking people who are living in their own homes, use of the telephone, particularly with people who do not know them, presents great difficulties.

A prototype was developed to investigate whether a goals-plans-script based user interface could be used in a 'real-time' dialogue situation with minimal risk of conversation breakdown and with adequate information transmission between participants. A script structure permits speech actions to be made available to participants in a context sensitive manner, while minimising the number and frequency of input actions needed to produce speech. The interface allowed the user to move easily up and down a goals-plans-scripts hierarchy. Part of the interface for the system is shown in Figure 1.

The system design assumed that the user would have at best a limited capacity for keyboard input (this would apply to nearly all persons with permanent speech loss). The prototype thus avoided presenting the user with large amounts of data to process.

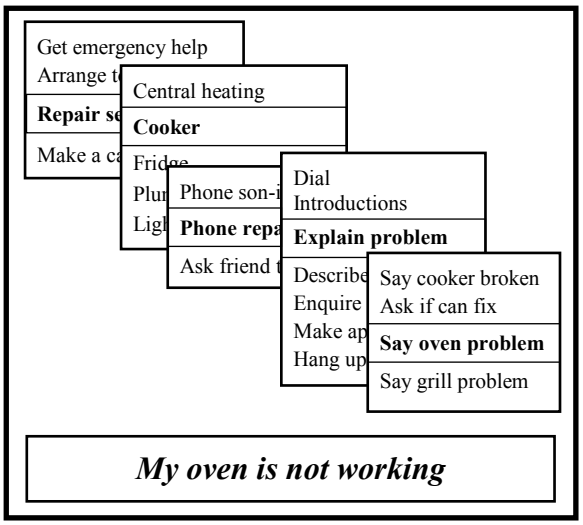

Fig. 1. Example of the interface design. The user moves down through a goal-plan-script hierarchy to finally select the wanted utterance.

Using the prototype, an experimental situation was devised whereby the experimenter attempted to complete a particular task over the telephone. The task was to telephone about getting a cooker (kitchen stove) repaired. Forty volunteer subjects were asked to play the role of an appliance repair person. They were spoken to over a telephone link by the experimenter. With one group of subjects the experimenter used his voice as normal, and with the second group he used only the script-based prototype, which produced speech through a speech synthesiser. In order to test fully the capabilities of using prestored speech, the prototype deliberately had no facility for the user to type in any new material. All contingencies had to be handled with the prestored scripts.

In all, 40 conversations were carried out with volunteer subjects. As expected, the computer-mediated conversations did take longer than the natural conversations, but the average was only 50% longer (Mean length of calls made with ordinary voice was 102 seconds [s.d. 14 seconds]. Mean length of calls made using the script-based speech output system was 150 seconds [s.d. 26 seconds]).

A small number of the subjects found the delays difficult to cope with. The mean delay time before the 'disabled' person spoke was 4.94 seconds [s.d. 2.19]. This is promising, in that it is approaching the 3 seconds 'awkwardness limen' identified by McLaughlin and Cody [6] as being the boundary between acceptable and unacceptable delays in conversation. In this situation, of course, the other speaker will also be aware that the augmented speaker will need a bit of extra time to speak in any case, but the very low speech rates currently obtainable do create real barriers to communication, and the fact that this prototype came close to a threshold established for natural speakers was encouraging.

Information transfer and effectiveness in accomplishing the given task were not significantly different between the two groups. In a questionnaire about the acceptability of the prototype, almost all subjects were positive about it. Some of the subjects had difficulty understanding the synthesised speech.

The evaluation established that a script-based approach to goal-oriented interactions shows promise as a means of helping a severely impaired non-speaking person to communicate effectively in these situations.

FUTURE WORK

The next phase of this research is to develop a complete communication system incorporating predicted sequences and scripts. The system will provide assistance with opening and closing sequences, giving feedback to another speaker, and accomplishing a wide range of everyday communicational tasks by means of scripts. A hierarchical structure based on goals and plans would provide a reasonable starting point for the design, but the project will also investigate more adaptive and flexible approaches to script presentation, whereby the user may move quickly from script to script without the need for traversing a hierarchical structure. The use of script segments and templates will also be investigated.

The project has the benefit of an advisory group composed entirely of non-speaking people, from whom many of the ideas for useful scripts and routines come. Speaking on the telephone has been highlighted by the group as a priority area for the prototype.

At the start of the project, a conversational model was developed, based on previous work by two of the partners, and also a review of the literature on modelling conversations. The literature review established that the situational nature of communication meant that the best approach to modelling a conversation was script-based, since this took into account the user's situation, and, in fact grows out of this.

These considerations led us to design a model of conversation, which is shown in increasing levels of complexity in Figures 2, 3 and 4. In Figure 2 the three basic stages of conversation are shown as Openers, Discussion, and Closers [5].

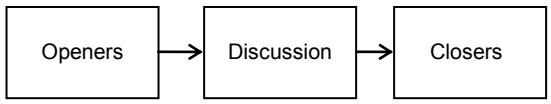

Fig. 2 A simple three-stage model of conversation

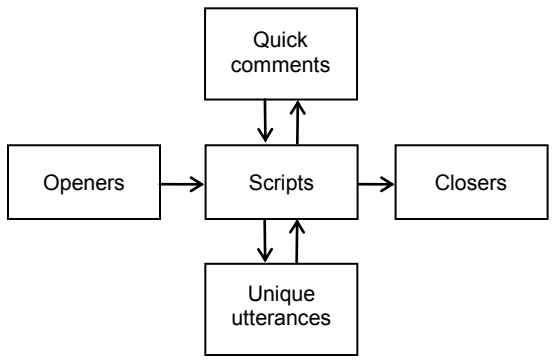

Fig. 3 Adding more structure to the Discussion stage

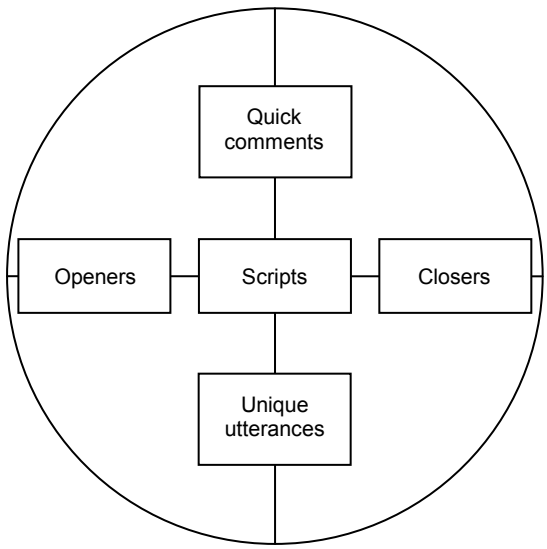

Fig. 4 Finally, version with each stage reachable directly from any other stage

The previous research on CHAT exploited the predictability of the opening and closing phases of an encounter [l]. A simple approximation for the Discussion phase, especially where a particular task is being accomplished, is a script facility, as described above. The CHAT research also established the usefulness of a ready supply of quick comments, particularly when another person is speaking [10]. So a script facility plus quick comments provides a better approximation to natural speech. Finally, a Unique Utterance facility will clearly always be necessary, though a time penalty will be paid when this is used. It must be a design aim of such a system to provide a good enough set of scripts so as to minimise the necessity for using the Unique Utterance facility. The model incorporating these features is shown in Figure 3.

Although many conversations will begin with Openers, and then proceed through Scripts to Closers, with the assistance of Quick Comments and Unique Utterances, the user must be free to be completely flexible in using the system. For instance, if a colleague is just rushing out the door, the user will want to go straight to a Closer and say something like 'Be seeing you'. Or if someone approaches with a quick question, the user will want a Quick Comment or two followed perhaps by a Closer, leaving out the usual preliminaries. The final diagram in Figure 4 shows each part of the model being reachable from each other part.

As can be seen from this model, the communication system user is free to move to any part of the conversation sequence, but the default sequence begins with openers, ends with closers, and has a middle section which consists of scripts and unique utterances when the user is speaking and readily available quick comments when another person is speaking. It is this model which will be used as the basis for the prototype system, now under development.

CONCLUSIONS

This investigation established that a goals-plans-scripts hierarchy was useful in enabling the user of a speech system with prestored utterances successfully to accomplish a useful task over the telephone. Although tested initially by a non-disabled person, the system performance showed enough significant improvement over existing communication systems to suggest that this method is worthy of further development. The next stage is being taken, with a communication system which is built up from a set of scripts, along with a quick comment and a unique utterance facility.

REFERENCES

[1] N. Alm, J.L. Arnott, A.F. Newell, "Prediction and conversational momentum in an augmentative communication system", *Communications of the ACM*, Vol. 35, No. 5, 1992, pp. 46-57.

[2] L. Broumley, J.L. Arnott, A.Y. Cairns, A.F. Newell, "TalksBack: An application of AI techniques to a communication prosthesis for the non-speaking", *Proceedings of the 9th European Conference on Artificial Intelligence,* Stockholm, Sweden, 6th-10th August, 1990, pp. 117-119.

[3] F. Dennis, A. Waller, A.Y. Cairns, J. Brodie, A.F. Newell, K. Morrison, "Investigating the use of a predictive conversation aid with dysphasic adults", *Abstracts of the Communication Matters Symposium,* 1993, p. 24.

[4] Don Johnston Incorporated, Talk:About™, Don Johnston Incorporated, 1000 N. Rand Road, Wauconda, Illinois, U.S.A.

[5] J. Laver, "Linguistic Routines and Politeness in Greeting and Parting", in Coulmas, F. (ed.) *Conversational Routine -- Explorations in Standardized Communication Situations and Pre-Patterned Speech,* The Hague: Mouton, 1981.

[6] M. McLaughlin and M. Cody, "Awkward silences: Behavioural antecedents and consequences of the conversational lapse", *Human Communication Research,* Vol. 8, 1982, pp. 299-316.

[7] A.F. Newell, J.L. Arnott, N. Alm, "Developments towards an integrated prosthesis for the non-vocal", *Proceedings of the 13th Annual Conference of the Rehabilitation Engineering Society of North America (RESNA '90),* Washington, D.C., U.S.A., 15th-20th June, 1990, pp. 97-98.

[8] R. Schank, R. Abelson, *Scripts, Plans, Goals, and Understanding,* New Jersey: Lawrence Erlbaum, 1977.

[9] A. Waller, "Providing narratives in an augmentative communication system", *Ph.D. Thesis*, University of Dundee, Dundee, Scotland, UK, 1992.

[10] V. Yngve, "On getting a word in edgewise", in *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, 1970, pp. 567-578.