Using NLP in the design of a conversation aid for non-speaking children

Portia File
School of Informatics
University of Abertay Dundee
Bell Street, Dundee DD1 1HG
Scotland, UK
p.file@tay.ac.uk

Leona Elder
School of Social and Health Sciences
University of Abertay Dundee
Bell Street, Dundee DD1 1HG
Scotland, UK
l.elder@tay.ac.uk

Abstract

It is difficult to develop an effective computer-based aid to enable children whose speech is hard to understand to participate in social conversations. Such children require a relatively simple device that will enable them to select and produce suitable content for a conversation in real time. We are investigating how an analysis of children's language can be used in the design of such a device. In particular, we believe that an analysis of the pragmatics of children's conversations will provide a basis for predicting from a corpus of plausible utterances those utterances that will be most useful for a particular child. Such an analysis may also be useful for subsequently predicting and locating which of these pre-loaded utterances is likely to be required next during a particular conversation. In addition, we are interested in using computer-based training for the aid. Given that all of the child's speech will be produced through the computer-based aid, it may be possible to provide intelligent interactive training.

1 Introduction

Children who cannot speak tend to be socially isolated from their peers. We are currently developing a computer-based device, called PICTALK, to support their casual conversation and hence their social development. Because the system is intended for children, we need careful control of the cognitive demands made on the system's users. We wish to investigate how the methods developed in NLP research can be used to improve the effectiveness of PICTALK in supporting the conversations of these users without increasing its complexity.

The PICTALK system (File et al., 1995a); (File et al., 1995b) is intended to support the casual conversation of people who can neither read nor speak. It is based on the TALK system (Todman, Alm, and Elder, 1994a) for people who can read but who are not able to speak. With PICTALK, pictures and labels are used to indicate the content of utterances that the user can choose to speak (using a speech synthesiser). Content utterances are stored within an organisational framework (provided for the user) that is designed to enable their prompt retrieval in 'real-time' conversations. The content available to a PICTALK user is very limited for two reasons. First, each PICTALK screen has few items or 'buttons' available because each picture takes up quite a bit of space. Further, the number of screens that can be accessed by a user using the organisational framework is restricted by the need to limit the complexity of the system to that suitable for someone who is either very young or who has learning difficulties. It is important to use any methods we can, e.g. predictive methods, that will reduce the cognitive load on the user or that will allow the user effective access to a larger set of utterances.

The content of informal conversational exchange is often subordinate to its social aspects. Therefore, our emphasis on casual conversation leads us to focus more on supporting those social aspects of conversation rather than on the delivery of information. Our principal goal is to allow the system's users and their conversational partners, working together, to have pleasant social interactions. To achieve this goal, we are examining the pragmatic structures of children's successful unaided conversations and would like to use the relationships between these structures to predict content.

2 The pragmatics of children's conversations

McTear (McTear, 1985) has examined the pragmatics of children's conversations. The main pragmatic structures he notes are: greetings, initiations, attention getting, attention directing, conversation repair e.g. repeating an utterance or requesting or responding to a need for clarification, and use of discourse connectors for topic shift or to continue the conversation after repair and to signal turntaking. Turntaking exchanges can be to initiate, respond, follow-up or conduct a simultaneous response with initiation (e.g. "Is it in the cupboard" in response to "where is it?"). Even young children use verbal and non-verbal means to accomplish these activities as well as changes in prosody and variations in politeness depending on the partner. We are interested in investigating ways in which we can assist children in carrying out these activities by using prediction within PICTALK's organisation structure to reduce the complexity of the process required for finding the next appropriate utterance.

3 PICTALK for Children

PICTALK has many features that support the above activities. These include the use of pre-loaded text, a menu controlled organisational structure that models conversation flow and additional items specifically designed to keep the flow of conversation going. There may be a potential for NLP to contribute to the enhancement of these and other aspects of the PICTALK system.

3.1 The role of pre-loaded utterances in PICTALK

The PICTALK system allows the user to pre-load potential conversation fragments that may be useful in some future conversational interaction. The substantive content of these fragments may be input with a particular interaction and a specific conversational partner in mind or it may be more general for use with any of a number of potential partners. This pre-loading process may be very slow but it can be carried out whenever the user has time to do it and under circumstances when there will certainly be much less time pressure than during actual conversation. The intention is that the user will be able to access these pre-loaded utterances quickly during the conversation. Such rapid access is likely to be very important in social conversation. Experience with the TALK system (Todman and Lewins, 1996) suggests that the rate of conversation has a strong positive relationship with the ratings of satisfaction made both by a TALK user and by her conversation partners.

Pre-loaded utterances currently have to be constructed by someone other than the PICTALK user Each of these utterances needs to be considered very carefully for several reasons. First, PICTALK holds very few (typically two dozen) utterances. As already mentioned, this is partly because few pictures can be accommodated on a fixed size of computer screen. Additionally, it is important that the cognitive demands made of PICTALK be limited and, therefore, the number of possible decisions that are required in order to select utterances must also be limited. Finally, PICTALK users are likely to have difficulty deciding how to make the conversation flow when the utterance they would like to use is not available.

3.2 Selecting pre-loaded utterances

It is generally considered highly desirable to allow people with disabilities to be as independent as possible. In this context, it would be desirable to allow PICTALK users to take more control over the content of their social interactions. Though at present utterances are developed and pre-loaded by someone other than the PICTALK user, the PICTALK system has a facility to allow the end user to select from a database of those available utterances and their associated pictures. Because the association between picture and utterances is imprecise, the user will need to experiment with the available items to see what speech is associated with each available picture before deciding whether to load an item into their PICTALK system. It would be helpful to offer first the items that are most likely to be appropriate. A wide range of conversation attributes could be used to predict the most appropriate utterance, e.g. the anticipated conversation partner, phrases selected for similar conversation partners, content related to the last phrase selected. Even a fairly small database of additional items with individually selected and stored attributes could offer benefits to these users.

3.3 Supporting variations in conversational style

At present it may be possible to support some variations in conversational style as a function of who the conversational partner is (e.g. peer, teacher). The simple solution is to address this problem when utterances are constructed for the user with a particular conversation partner in mind or when utterances are selected during a conversation. For example, in PICTALK variations in style can be expressed in the

four utterances that are available to open a conversation (e.g. *hiya; howdy; hello*) and the four other utterances that are available to close a conversation (*see you later; goodbye; see ya*). It may be possible to provide more general support for variation in style either by picking up the cue to style from the opening utterance selected and then using the corresponding variation of each subsequent utterance selected or by offering suitable prediction when the user selects material from a database of utterances in preparation for an interaction with a particular class of conversation partner e.g. polite style for adult, more informal style for classmate.

3.4 Accessing material through the organisational structure

Most of the utterances available to the PICTALK user are organised in a shallow menu structure. There is a hello menu button to give access to greeting utterances, a goodbye menu button to give access to closing utterances and most of the remaining content is accessed through a set of 3 intersecting perspectives, namely, person with 2 values: me and you; tense with 2 values: past and present/future and affect with two values: happy and sad. By selecting one each of these three perspectives, the user gets access to pre-loaded content appropriate to that combined perspective e.g. content on something I like doing (me/present/happy). These perspectives are designed to support the flow of conversation and require only one button press to move from phrases about what I disliked (me/past/sad) to phrases about what you disliked (you/past/sad) or from phrases about what I liked (me/past/happy) to phrases about what I would like (me/present/happy). With similar features, TALK has been shown to support its users in initiating conversation topics and in turn-taking in the Question-Answer format (Todman et al., 1994b). Unfortunately, our experience with PICTALK users suggests that this menu structure is very difficult for them to understand. It may, however, be possible to support the PICTALK user by predicting and suggesting suitable utterances during a conversation.

From McTear's (McTear, 1985) work, it seems plausible that children may be able to recognise a suitable utterance (identified by its associated picture) as appropriate if it were suggested, even if they are unable to recall and locate it. Though it would be difficult to implement, a modestly effective prediction system could reduce the cognitive load on the PICTALK user. Such a prediction system may be easier for children than for adults. Some theorists, notably, of course, Piaget (Piaget, 1959), have suggested that children are more egocentric than adults in all aspects of social relationships including conversations, by which it is usually taken to mean that they respond less to the listener's perspective. While later work has suggested that the level of egocentrism which is in general exhibited by young children may not be as great as suggested by Piaget e.g. (Selman, 1980) nevertheless the ability to make this type of social adaptation is not fully developed for some years and therefore children may be less responsive to their partners (McTear, 1985). This may have the fortuitous consequence that the child's next utterance depends more on the child's last utterance, which is known by the system, rather than on the partner's last utterance, which is not known.

3.5 Keeping the conversation flowing

PICTALK provides a few other utterances that are outside of the menu structure and are always available. The utterances are intended to support the goal of maintaining conversational flow when a suitable specific response to something a partner says is not available and fulfil general pragmatic functions such as initiation to get attention, e.g. "Hey!", repair when something has gone wrong, e.g. "OOPS" or "What?" or "I don't have anything to say to that" and discourse connection to signal a topic shift, e.g. "Now, right". Utterances can also be available support the speaking conversation partner. In using the TALK system, we have found that keeping the partner informed in some way is reassuring and helpful. They need to know, for instance, that a longer than usual silence is not signalling the termination of the interaction and this can be achieved through the inclusion of such utterances as "hang on. I need to find what I want to say". It may be that some of these may be automated. For example, it might be possible for the PICTALK system to recognise that its user is searching for something to say when consecutive menu presses are made or even to recognise that a topic shift is being initiated when the utterance just selected differs greatly in semantic content from the previous utterance.

3.6 Training

Children who have no speech or whose speech is so impaired that they cannot be readily understood may not have had opportunities to develop a familiarity with the structure and pragmatics of conversational interactions nor the skills required to maintain them. They also need considerable training to use the PICTALK system effectively. Therefore, we are developing a training system to help such children develop their conversation skills with PICTALK. In

the training system, the computer, using one speech synthesised voice, converses with the child, using another speech synthesised voice. At present the script for the computer is entirely pre-programmed. It might be possible to introduce intelligence into the script to allow the computer to respond appropriately to a wider range of user's contributions. Such a system should make it easier to retain the child's interest during the rather long training period. Because each of the child's responses is known to the system and is predictably limited to those available in PICTALK, it might be possible for the computer to converse fairly naturally with the child.

4 Current research

At present we are carrying out a preliminary investigation of the use of NLP to support the use of PICTALK by children working in structured settings to achieve the following:

- Analyse the pragmatic features of children's conversation in a particular setting e.g. a structured news setting.
- Analyse the variability in the features used by different children in the setting and eventually by the same children in different settings. Try to associate different conversational styles with different utterances for later use in predicting which utterances will be appropriate.
- Use these analyses to develop a corpus of plausible utterances together with an indication of which features are likely to be most helpful for children with different conversational styles or in different conversational settings.
- Use the utterance features to predict which utterances will be useful in the chosen setting and to help individual children to select utterances from the corpus to include in PICTALK for use in the setting.
- Use the analysis to see if there is any scope for the prediction or automatic insertion of utterances during a conversation.
- Use artificial intelligence to script PICTALK training conversations.

References

Portia File, John Todman, Norman Alm, Leona Elder, and Hilary Smith. 1995a. PICTALK; A conversation aid for nonspeaking, nonreading people. *Journal of Rehabilitation Sciences*, 8:47.

Portia File, John Todman, Norman Alm, Leona Elder, and Hilary Smith. 1995b. Adapting a Conversation Aid for Non-Speaking People (TALK) to the Needs of Non- Reading, Non-Speaking People (PICTALK). In *Proceedings European Conference on the Advancement of Rehabilitation Technology*, pages 158–160.

Michael McTear. 1985. *Children's Conversations.* Blackwell, Oxford, England.

Jean Piaget. 1959. *The Language and Thought of the Child*, 3rd edition. Routledge, London, England.

R.L. Selman. 1980. *The Growth of Interpersonal Understanding*, Academic Press, New York.

John Todman, Norman Alm, and Leona Elder. 1994a. Computer-aided conversation: a prototype system for non-speaking people with physical disabilities. *Applied Psycholinguistics*, 15:45–73.

John Todman, Leona Elder, Norman Alm, and Portia File. 1994b. Sequential Dependencies in Computer-Aided Conversation. *Journal of Pragmatics*, 21:141–169.

John Todman, Elizabeth Lewins, Portia File, Norman Alm, and Leona Elder. 1995. Use of a Communication Aid (TALK) by a Non-speaking Person with Cerebral Palsy, *Communication Matters*, 9:18–25.

John Todman and Elizabeth Lewins. 1996. Use of an AAC system for casual conversation. In *Proceedings 7th Biennial Conference of the International Society for Augmentative and Alternative Communication*, pages 167–168.