Pragmatics and AAC approaches to conversational goals

John Todman
Department of Psychology
University of Dundee
Dundee, DD1 4HN
Scotland, UK
j.todman@dundee.ac.uk

Norman Alm
Applied Computing Division
University of Dundee
Dundee, DD1 4HN
Scotland, UK
nalm@mic.dundee.ac.uk

Abstract

The modelling of pragmatic features of natural conversation to help AAC users achieve a range of social conversational goals is considered in relation to the development of an AAC system based on text pre-storage and retrieval. Problems facing designers of AAC systems are highlighted in the hope that insights from NLP may contribute to solutions.

1 Conversational goals

People have a variety of implicit and explicit goals when engaging in conversation, among which a broad distinction between social goals and goals concerned with "getting things done" can be discerned. For example, Cheepen's (Cheepen, 1988) distinction between transactional goals and interactional goals has considerable overlap with Hobbs and Evans' (Hobbs and Evans, 1980) ideational and interpersonal categories.

On one hand, transactional and ideational goals are broadly applicable to contexts in which something external to the conversation is "being done" (e.g. evolving plans, engaging in a task) and, on the other hand, interpersonal and interactional goals tend to predominate when the focus is on social aspects of the conversation itself. We will use the labels "transactional" and "social" to refer to this broad distinction.

Goals concerned with getting things done in the world, the transactional goals, may range from getting a snack prepared to your liking, through planning an outing, to gaining a qualification. In general, the successful achievement of such goals will rely heavily on the accurate transmission of information during the relevant communication episodes. For these "message oriented" goals, as Brown and Yule (Brown and Yule, 1983) describe them, precision of meaning in the content of the conversation is likely to be crucial.

For social goals, on the other hand, described by Brown and Yule as "listener oriented", precision of conversational content may sometimes be less important than aspects of delivery, especially timing. As with transactional goals, social goals may range from the immediate, such as enjoyment of a social interaction or making a favourable impression, to longer term goals such as the development of relationships or self esteem.

Short and long-term goals may, of course, both be active with respect to the same conversation, with short-term goals contributing to longer-term goals, which, in turn, contribute to very long-term goals such as "quality of life" and "self-fulfilment". Similarly, conversations are not necessarily exclusively concerned with either social goals or transactional goals. For example, a casual chat may contain transactional components such as arranging future joint activities. Conversely, a basically transactional conversation with a shop assistant may include some social chat. Nonetheless, the broad distinction between conversations motivated primarily by social goals and those motivated primarily by transactional goals can be sustained.

2 AAC design approaches

For people who are unable to speak, most high-tech. AAC systems aim to help them to communicate using synthesized speech. This can be approached either by means of pre-storing complete phrases ready for retrieval and output in a subsequent interaction or by constructing phrases at the time they are required during an interaction. Most current AAC systems have been designed using a predominantly phrase-construction approach in order to maximize the flexibility of speech output. In these phrase-construction systems it is common for some sort of

prediction to be incorporated to ease and speed the task of entering content and for provision to be made for storage of a few frequently used messages.

It is implicit in this approach that the precision of expression attainable with flexible phrase construction at the time a thought occurs is the paramount consideration. Pre-storage of phrases is, in this view, seen as suitable only for passing simple, frequently used messages, for very routine exchanges such as greetings and good-byes and for delivery of non-interactional monologue, as in giving an address. Pre-stored phrases are not generally considered suitable for the conduct of free-flowing social conversation, where it is assumed that subtle nuances of meaning need to be constructed as the conversation proceeds in directions that could not have been foreseen.

This view of social conversation may, however, be misleading in its disregard of the range of goals that motivate such interactions. If the main point of a conversation is simply to enjoy the social interaction, it has to be asked whether the conversation will be more enjoyable with long pauses while each "ideal" utterance is constructed or with roughly "appropriate" utterances that are delivered without long pauses preceding them. Similarly, it may be asked which of these scenarios is more likely to result in an attribution of, say, competence or, in the longer term, to have a positive effect on an AAC user's self-esteem, status or independence.

Certainly, users of high tech. AAC systems complain more about the slowness of their speech output than about any restrictions limiting the precision with which their thoughts can be expressed. This is not just because most users are working with phrase-construction systems, which do not impose restrictions on precision (except those due to time constraints). In a recent study, Todman and Lewins (Todman and Lewins, 1996) trained a non-speaking person to use a text storage and retrieval AAC device to engage in free-flowing conversation at a rate of output (counting search times) almost 10 times faster than her output using her usual phrase-construction system. The critical finding was that the faster her speech output in a particular conversation, the more pleasurable that conversation was rated by both the user and her conversational partners. This is consistent with what is known of the disruption caused by long pauses in natural conversation (McLaughlin and Cody, 1982).

Another feature of natural conversation that has an important bearing on which general approach to the design of AAC systems (i.e. phrase-construction or phrase-storage) is more likely to help users to achieve their conversational goals is the imprecision of much conversational content. Wyer and Gruenfeld (Wyer and Gruenfeld , 1995) argue, for example, that "if-then" production rules suffice to produce acceptable routine responses to many of the things a conversational partner may say, and Langer (Langer, 1978) provides evidence that carefully considered, precise responses are the exception rather than the rule. Responses need to be "appropriate" but they do not need to be "ideal" or precise to meet participants' goals in much social conversation.

People who are unable to speak tend to be socially isolated. Social goals are therefore likely to be particularly salient for them. Therefore, the development of AAC devices capable of supporting a reasonable approximation to natural social conversation should be a high priority for AAC designers. However, casual social conversation, with it's free-ranging content and its dependence on speed of responding, presents a considerable challenge. It seems reasonably clear that AAC systems based on phrase-construction are unable to meet some of the social goals that users are likely to have. Although there are substantial difficulties to be overcome in the development of phrase-storage systems, their potential for outputting responses relatively quickly suggests that they may be capable of meeting at least some of the more immediate social goals of AAC users. When the aim is to develop an aid for social conversation, there seem to be good reasons for adopting a basically phrase-storage approach.

A predominantly phrase-storage system will need features that deal with the impossibility of anticipating precisely what phrases will be needed in a subsequent conversation and with the added difficulty of locating stored phrases for output (Light et al. , 1990). Furthermore, to produce an effective AAC system, even for just social conversation, phrase-construction features will need to be incorporated within the basically phrase-storage system. A reasonable way of approaching the design issues is to consider what pragmatic features of natural conversation seem to support various goals that the participants may have, with a view to modelling such features in the AAC system.

3 Pragmatic features of natural conversation

The need for responses generally to be *appropriate* and *fast* rather than ideal and slow has already been discussed. Sometimes, however, it will be necessary to generate a *unique* response. Recycled, imprecise responses will not do when specific information is

requested. Again, participants have to be able to *cope with the unexpected.* Generating a unique response may be one way of doing this but it is by no means the only strategy that is effective in natural conversation. People frequently respond in a rather general way, make "hedging" comments to gain time or explicitly defer discussion of the topic to a later occasion.

The general requirement for speed of output bears on some other features of natural conversation that are speed-dependent to varying degrees. Co-operative efforts are required to *maintain flow*, with the orderly development of topic being managed by means of topic shifts that are small enough to maintain continuity, with occasional larger shifts to establish new directions. Another way in which co-operation is evident in natural social conversation is in the *sharing of control* of topic direction. Indeed, Cheepen (Cheepen, 1988) has argued that this is a defining feature of interactional (i.e. social) conversation. Participants in natural social conversation further demonstrate their co-operative involvement with frequent positive *feedback* while a partner is making an extended contribution to the conversation and by means of *repair* strategies when things go wrong, threatening breakdown of the conversation.

These pragmatic aspects of natural conversation contribute in different degrees to the various short and long-term social goals of participants. It is important to incorporate features within an AAC design which, by modelling such aspects of natural conversation, will help users to pursue their social goals more effectively than has been possible so far. The first requirement is to develop a basically phrase-storage system that emphasizes speed of output and models other speed-dependent features of natural conversation (i.e. maintenance of flow, share in control, feedback, repair). This will be particularly important for meeting immediate social goals, such as enjoyment of the interaction and creation of a favourable impression, which is essential if the user is to remain motivated to use the aid to have social conversation.

The next requirement is for design features that model those aspects of natural conversation that are particularly content-dependent (i.e. uniqueness, appropriateness, coping with the unexpected). These content-dependent features are needed, together with the speed-dependent features, to meet longer-term goals such as those concerned with the development of relationships, participation in activities, status, self-esteem and independence. Although content-dependent features of conversation can be modelled to some extent within a phrase-storage approach, this will have to be supplemented by a phrase-construction component.

The proposed relationships between AAC design approaches, pragmatic aspects of natural conversation and short and long-term conversational goals are illustrated in Figure 1. As an example of how the design of an AAC system can be guided by the joint consideration of these variables, the development of a conversation aid known as "TALK" (Todman, Alm, and Elder, 1994) will be described.

4 The TALK system

The TALK system was developed in order to experiment with a number of ways of achieving conversational goals more easily using an AAC system. From the development of a number of previous prototype systems, some important lessons had been learned. Ways of providing an AAC user with rapid and effective speech acts for opening and closing a conversation have been devised (Alm, Newell, and Arnott, 1987). The importance of backchannelling in communication has been recognised and a way of providing quick-fire comments has also been developed (Alm, Arnott, and Newell, 1992a). Several different approaches had been taken to assisting an AAC user to handle the central part of a conversation. A text database of conversational material has shown the effectiveness of labelling stored items with their pragmatic as well as their semantic aspects, but proved difficult to use in practice without time-consuming construction of retrieval requests (Alm, Arnott, and Newell, 1989). Attempts to provide a degree of prediction of conversational material have included using narrative sequences (Alm, Arnott, and Newell, 1992b); (Waller, 1992), taking the participants' personal characteristics and interests into account (Broumley et al., 1990), and using fuzzy information retrieval techniques (Alm, Nicol, and Newell, 1993).

From all this work, together with the body of research into discourse and conversation, it was apparent that a simple, though partial, model of conversational interaction could be constructed. Given the incomplete nature of knowledge about conversational structure, such a model would of course be at best a good guess, but it could help clarify our thinking about how to build more effective systems for helping AAC users to accomplish conversational goals. The suggested model we have arrived at is shown in Figure 2.

The opening and closing of a conversation can be done according to a fairly well set out routine. In the central part of the conversation, the conversa-

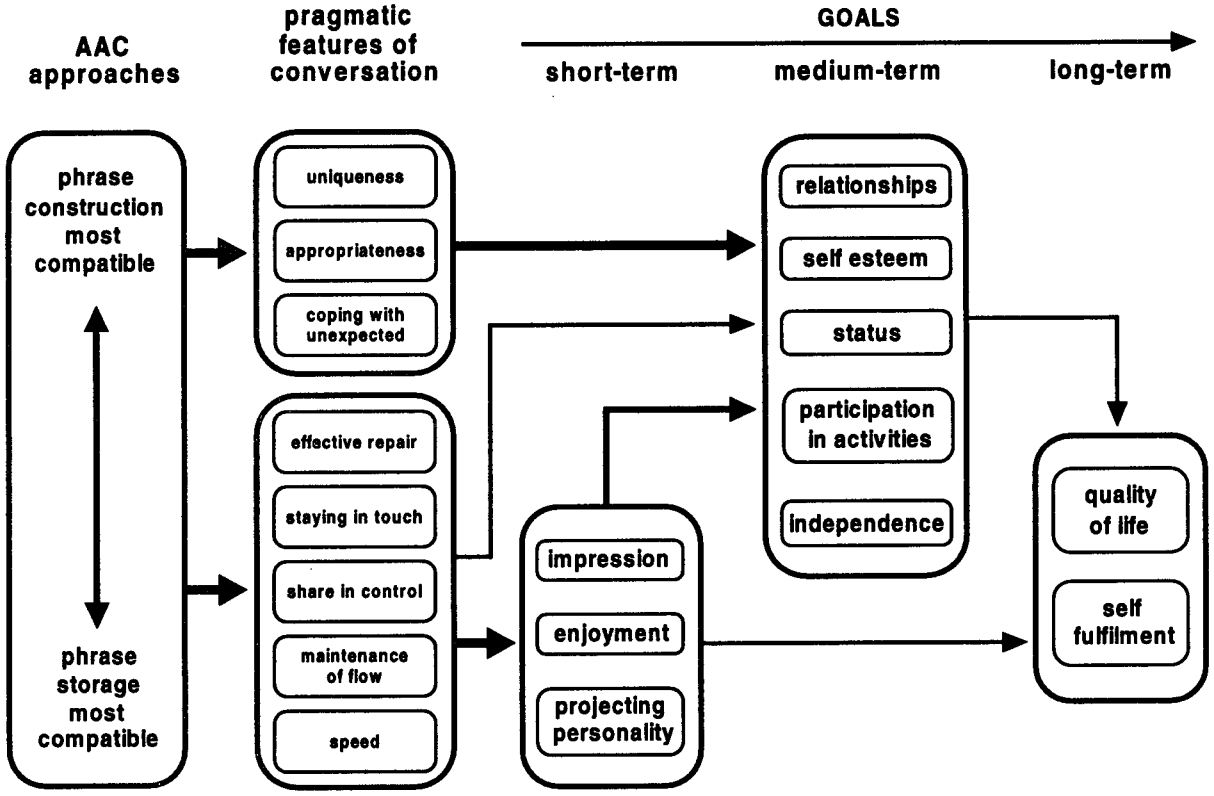

Figure 1: A model linking AAC design approaches, pragmatic features of conversation and user goals

tionalist is either speaking or listening. This is of course an oversimplification, since overlapping talk is in fact the norm. More accurately, we could say that the conversationalist is either leading the direction of the conversation or is following the other person's lead. When another is speaking, the conversationalist needs a quickly available supply of feedback remarks to express their reactions. When the conversational lead is being taken, they will need a way of speaking on a topic, and changing topics.

The TALK system incorporated features from an earlier prototype to handle opening and closing a conversation and giving quick feedback easily (Alm, Arnott, and Newell, 1992a). With the TALK prototype two important new features were introduced: a method of dealing with topic shift, and the inclusion of another important category of speech act: context-sensitive comments.

The easy handling of topic shift is a major problem for users of an AAC system. Given the slowness of operation, and the potential complexity of a system which could handle large amounts of text, some sort of predictive or assistive mechanism will be necessary to make topic shifting a realistic possibility. Such a capability would be important for achieving appropriate speed and maintenance of flow. Also, topic shift is a key method of sharing in the control of the conversational direction.

The TALK system handled topic shift by providing users with three sets of conversational perspectives. The user could shift these perspectives with one activation of an on-screen button. The perspectives were:
Person: Me, You
Time: Past, Present, Future
Orientation: Where, What, How, When, Who, Why.

By altering one of these perspectives the user called up a new set of candidate texts for speaking reflecting that perspective. For example, to shift from a screen displaying content related to how things had occured in the user's past (me/how/past) to a screen containing content related to how things

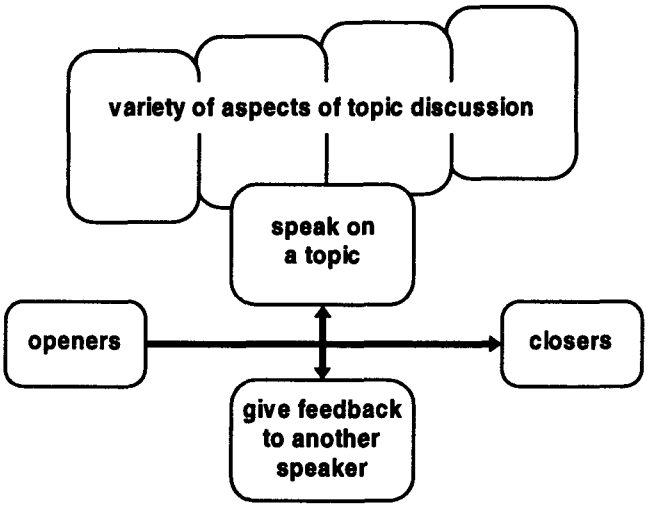

Figure 2: A simple and incomplete model of conversational interaction

occurred in the partner's past (you/how/past), the "You" button would be activated. Making a series of these selections took the conversation through the step-like progression from one topic to another which coherent conversations require (Button and Casey, 1984).

The TALK user had available a set of on-screen buttons with which to produce quick-fire responses to what another person was saying. These assisted in maintaining speed, giving appropriate feedback to another speaker and being able to effect repairs. The responses available were:
Acknowledge
Say yes, Say no, Say don't know
Agree, Disagree
Evaluate good, Evaluate bad
Interrupt, Say thanks, Ask for expansion
Say wait a minute (stall for time)
Say a mistake was made in speaking.

In order to maximise the speed of response, these phrases were spoken when the button was activated, without recourse to a menu of possible choices. A random variation was built in to avoid too much repetition of exactly the same words. For example, a set of alternative "acknowledge" responses might be "Uh-huh", "Yeah", "I see", "Yeah, yeah", "Yeah, uh huh". This was in keeping with the principle that in this case speed of response was the key issue, and if the phrase was not exactly what was required, an approximation to the wording needed would in any case be sufficient.

From early trials of the TALK system, the need for another category of speech act emerged. Often it is important to provide a comment which does need to be selected from a menu of possibilities, because its use is dependent on the context. Also there are reusable phrases which serve to move the conversation forward and which, although suitable as responses to many different things a partner might say, need to be selected specifically. These phrases we called "context-sensitive comments" (Todman, Alm, and Elder, 1994). We experimented with a number of different types of comment and, though the set that finally went into TALK was by no means a definitive one, the list of comment categories given below, with an example of each, was found to be useful (Todman and Morrison, 1995) and added to the flexibility of the system:
Aphorism (e.g. "That's how life goes sometimes")
Expression of sympathy (e.g. "Sorry to hear that")
Hedge (e.g. "I don't really remember")
Apology (e.g. "Sorry, I didn't think of that")
Question ("How about you?")
Specific feedback ("That's really interesting")

These context-sensitive comments, like the quick-fire phrases, helped with speed, maintenance of flow, and having a share of the control of the conversation. They also provided a better way of responding appropriately to the unexpected than the more general-purpose quick-fire remarks.

Some initial testing of TALK was performed in which only pre-stored text was used, in order to examine the limits of speaking entirely with pre-stored material. A facility for adding unique text for speaking during a conversation was then added. The use of this feature, of course, involves the user in a significant time penalty.

In trials, the TALK system has shown that incorporating the modelling of pragmatic features of conversation can produce improved results in computer-aided communication. Significant increases in speed are possible. One physically disabled non-speaking person using TALK achieved a speaking rate of about 67 words per minute (Todman et al. , 1995). This represents a considerable increase on the 2-10 words per minute which is the current norm (Beukelman and Mirenda, 1992).

In a study analysing the quality of the content of TALK-aided conversations compared with conversations on the same topic carried out by natural speakers, the content of the computer-aided conversations was rated significantly higher than that of the unaided samples ($p < .001$) (Todman, Elder, and Alm, 1995). This was an encouraging finding indicating that using pre-stored material could actually enhance the perceived quality of conversational content.

Currently two AAC users are taking part in a long-term evaluation of the usefulness of the TALK prototype in their daily conversations, and a third AAC user is evaluating a version of TALK which has been adapted for people with limited literacy skills. From these informal evaluations, the users report that using the prototype has given them conversational opportunities they would not otherwise have had. Two of the users have given a number of public lectures, using the TALK system to deliver the lecture and deal with the following question and answer sessions (Grant, 1995); (McGregor, 1995); (Todman and Grant, 1996). Such applications of the system have a clear relationship to the communicational goals of mutual enjoyment and enhancement of the perceived status of the speaker. These longitudinal long-term studies continue.

A number of improvements suggest themselves to further enhance the usability of systems such as TALK. The prototype models several aspects of the pragmatics of conversation, but no doubt there are other aspects which could be helpfully incorporated. As can be seen in Figure 3, the interface is quite complex at present. Introducing predictive features could help to simplify the control task for the user. It may be that work currently underway in the field of natural language processing can be of assistance in suggesting ways to accomplish this task.

References

Norman Alm, Alan Newell, John L. Arnott. 1987. A communication aid which models conversational patterns. In *Proceedings of the 10th Annual Conference of the Rehabilitation Engineers Society of North America*, pages 127–129.

Norman Alm, John L. Arnott, Alan Newell. 1989. Database design for storing and accessing personal conversational material. In *Proceedings of the 12th Annual Conference of the Rehabilitation Engineers Society of North America*, pages 147–148.

Norman Alm, John L. Arnott, Alan Newell. 1992a. Prediction and conversational momentum in an augmentative communication system. *Communications of the ACM*, 35:46-57.

Norman Alm, John L. Arnott, Alan Newell. 1992b. Evaluation of a text-based communication system for increasing conversational participation and control. In *Proceedings of the 15th Rehabilitation Engineers Society of North America International Conference*, pages 366–368.

Norman Alm, M. Nicol, Alan Newell. 1993. The application of fuzzy set theory to the storage and retrieval of conversational texts in an augmentative communication system. *Proceedings of RESNA '93*, pages 127–129.

David Beukelman and Pat Mirenda. 1992. *Augmentative and Alternative Communication: Management of Severe Communication Disorders in Children and Adults.* Paul Brookes Publishing Company, Baltimore, Maryland.

Liz Broumley, John L. Arnott, Alaistair Y Cairns, and Alan F. Newell. 1990. TalksBack: An application of artificial intelligence techniques to a communication prosthesis for non-speaking people. In *Proceedings of the European Conference on Artificial Intelligence*, pages 117– 119.

G. Brown and G. Yule. 1983. *Teaching the Spoken Language.* Cambridge University Press, Cambridge.

G. Button and N. Casey. 1984. Generating topic: the use of topic-initial elicitors. In J. Atkins and J. Heritage, editors, *Structures of Social Action-Studies in Conversation Analysis.* Cambridge University Press, London.

C. Cheepen. 1988. *The Predictability of Informal Conversation.* Pinter Publishers, London.

Sylvia Grant. 1995. Using TALK. *Communication Matters*, 9:22–25.

J. R. Hobbs and D. A. Evans. 1980. Conversation as planned behavior. *Cognitive Science*, 4:349–377.

E J. Langer. 1978. Rethinking the role of thought in social interaction. In J. Harvey, W. Ickes, and R. Kidd, editors, *New directions in Attribution Research*, 2:35–58, Lawrence Erlbaum, Hillsdale, New Jersey.

J. Light, P. Lindsey, L. Siegel, L. and P. Parnes. 1990. The effects of message encoding techniques on recall by literate adults using AAC systems. *Augmentative and Alternative Communication*, 6:184–201.

Alan McGregor. 1995. A voice for the future. In *Proceedings of the European Conference on the Advancement of Rehabilitation Technology (ECART '95)*, pages 127–129, Lisbon, Portugal, October.

M. L. McLaughlin and M. J. Cody. 1982. Awkward silences: Behavioral antecedents and consequences of the conversational lapse. *Human Communication Research*, 8:299–316.

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

John Todman, Leona Elder, and Norman Alm. 1995. An evaluation of the content of computer-aided conversations. *Augmentative and Alternative Communication*, 11:229–234.

John Todman and Sylvia Grant. 1996. Conversation using an AAC system based on pre-stored text. In *Proceedings of the 8th Biennial Conference of the International Society for Augmentative and Alternative Communication*, pages 169–170, Vancouver, Canada, August.

John Todman and Elizabeth Lewins. 1996. Use of an AAC system for casual conversation by a non-vocal person with cerebral palsy. In *Proceedings of the 7th Biennial Conference of the International Society for Augmentative and Alternative Communication*, pages 167–168, Vancouver, Canada. August .

John Todman, Elizabeth Lewins, Norman Alm, and Leona Elder. 1995. Use of a communication aid (TALK) by a non-speaking person with cerebral palsy. *Communication Matters*, 9:18–21.

John Todman and Zara Morrison. 1995. The use of non-specific comments in a conversation aid for non-speaking people. *International Journal of Rehabilitation Research*, 18:68–69.

Annalu Waller. 1992 Providing Narratives in an Augmentative Communication System. Ph.D. Thesis. University of Dundee, Scotland.

R.S. Wyer and D.H. Gruenfeld 1995. Information processing in interpersonal communication. In D.E. Hewes, editor, *The Cognitive Bases of Interpersonal Communication*, pages 19-47, Lawrence Erlbaum, Hillsdale, N.J.

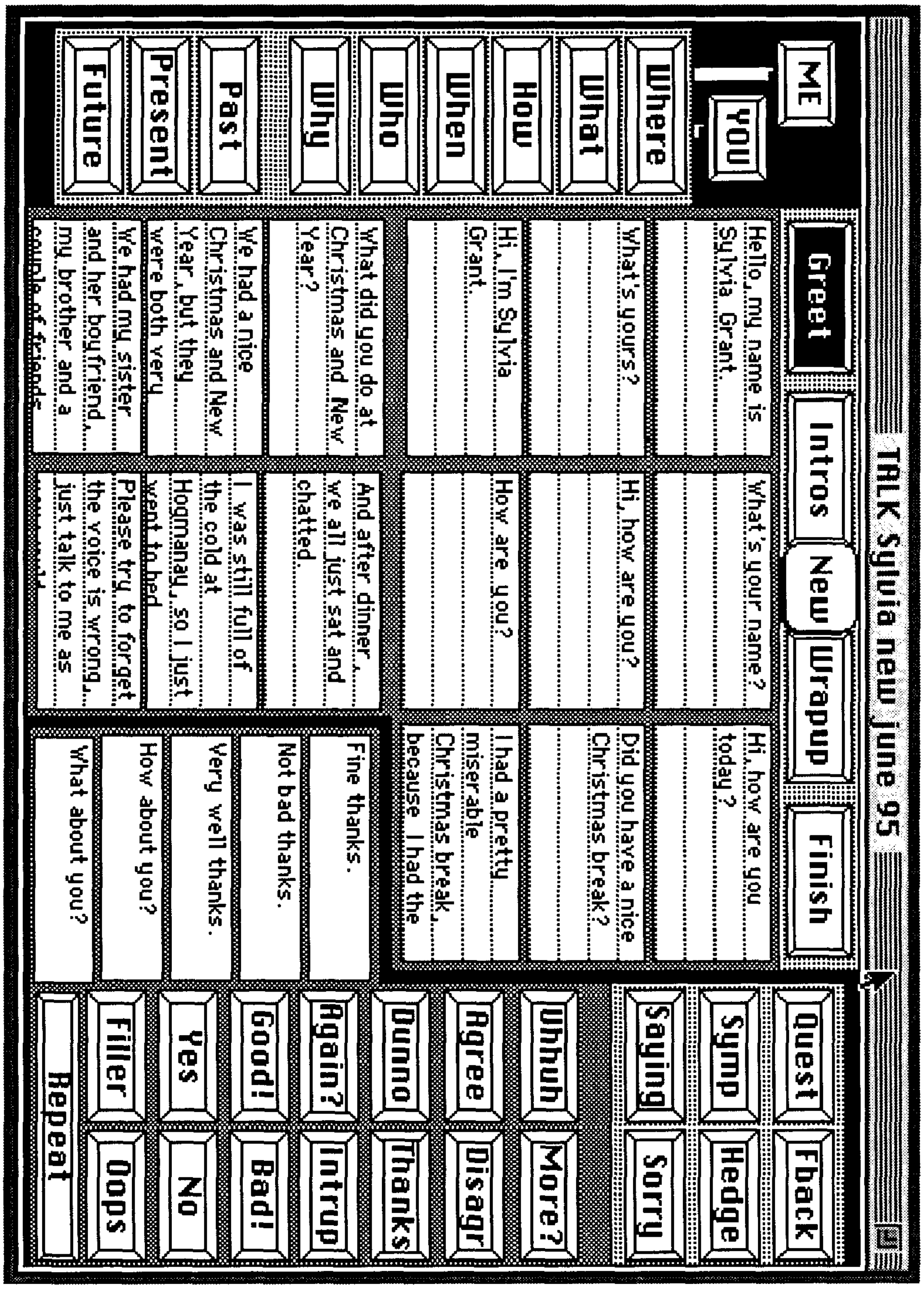

Figure 3: The TALK interface