Towards the improvement of augmentative and alternative communication through the modelling of conversation

Arnott, John L.; Alm, Norman

Published in:
Computer Speech and Language

DOI:
10.1016/j.csl.2012.10.008

Publication date:
2013

Towards the improvement of Augmentative and Alternative Communication through the modelling of conversation. J.L.Arnott and N.Alm, *Computer Speech & Language*, Vol.27, No.6, Sept. 2013, pp.1194-1211.

Towards the improvement of Augmentative and Alternative Communication through the modelling of conversation

John L. ARNOTT * and Norman ALM

School of Computing, University of Dundee, Dundee DD1 4HN, U.K.

{ jarnott / nalm } at *computing.dundee.ac.uk*

Computer Speech & Language is at: http://www.journals.elsevier.com/computer-speech-and-language/

ISSN: 0885-2308

URL: The DOI bookmark for the article is: http://dx.doi.org/10.1016/j.csl.2012.10.008

Abstract: Non-speaking people who use Augmentative and Alternative Communication (AAC) systems typically have low rates of communication which reduces their ability to interact with others. Research and development continues in the quest to improve the effectiveness of AAC systems in terms of communication rate and impact. One strategy involves making the basic unit of communication an entire utterance, and designing the AAC system to make the storage, retrieval and production of utterances as easy and efficient as possible. Some approaches take this further and include texts, narratives and/or multimedia material for use in conversation. AAC systems operating in such a manner require a structure for containing and managing conversational material and supporting the production of output during conversation. Ideally such a structure should be modelled on the way actual conversations proceed. A number of partial models for this have been presented thus far. These are reviewed in the paper and an integrated model is then proposed that includes both the structure of a conversation and the way in which an AAC system might produce conversational output (e.g. utterances, texts, multimedia items or combinations of these). Modelling the process in this way gives a structure with which an AAC system can organize the support and guidance that it offers to the person using the system. The paper concludes with consideration of three areas of development for further investigation.

Keywords: AAC, augmentative communication, conversation, modelling, non-speaking, scaffolding.

*** Corresponding author:** Prof. John L. Arnott, School of Computing,
University of Dundee, DUNDEE DD1 4HN, Scotland, UK.

1. Introduction

Inter-personal interaction is more difficult for people who use Augmentative and Alternative Communication (AAC) than for natural speakers. People using AAC typically have very low rates of communication because of their disabilities, so communication acts produced during real-time interaction tend to be brief and delayed. It is reported that conversation can proceed at rates of 150 to 250 words per minute for natural speakers who have no disabilities, whereas people who need to use AAC are limited to much lower rates, for most less than 15 words per minute (Foulds, 1980; Beukelman and Mirenda, 2005). A typical AAC interaction will involve a disabled non-speaking person using an AAC system to communicate with a speaking person. Because of the slowness of the interaction, it is very difficult for people to relate narratives or hold forth for any extended period of time if they are dependent on AAC for communication. This tends to reduce the spontaneity and vivacity of interactions and to cause them to be of relatively short duration. It also influences attitudes and expectations that speaking people have regarding interactions involving people who use AAC (McCarthy and Light, 2005).

Researchers have investigated many avenues for increasing the interactive communication rate and responsiveness possible in AAC (Beukelman and Mirenda, 2005). Some notable success has been achieved in this area, although communication rates still fall short of those achieved by speaking people who do not need to use AAC. The principal advance has been achieved through the use of whole utterances (e.g. phrases and sentences) pre-stored in the AAC system and retrieved later at the point-of-use when they are required during live conversation (Baker, 1982; Waller et al., 1990; Vanderheyden et al., 1996; Langer and Hickey, 1998; Alm and Arnott, 1998; Higginbotham et al., 1999; Todman and Alm, 2003; Hine et al., 2003; Todman et al., 2008; Dempster et al., 2010a). This was aimed at assisting non-speaking people to efficiently and effectively retrieve and use pre-stored text sequences (e.g. stories, personal biographical material, anecdotes, family news) as well as phatic remarks such as greetings, responses and farewells, and to do so in a conversational situation. Without this approach, it is necessary for utterances to be constructed in real time during conversation, which is very slow and limits greatly the flow and liveliness of a conversation. While it is important for on-line construction of utterances to be possible in an AAC setting, so that unique statements can be made by anyone using AAC, it is the use of pre-stored or automatically generated utterances that has brought most of the increased pace and responsiveness to AAC conversation that has so far been achieved. Output of such utterances to conversation partners is achieved typically via synthetic speech; an utterance-based AAC system usually has a speech output modality, in which case it can be described as a VOCA (voice-output communication aid) or an SGD (speech generating device). Other output modalities are also possible; visual display of messages to communication partners can be used, for example, and multimedia material (images, audio and video) can be included.

As the AAC system needs facilities for on-line construction or modification of output items as well as search and retrieval methods for accessing stored items, there is a need for the system to know where it is in a conversation, how the various stages of a conversation fit together, and the sequence of events that are likely to occur during a conversation. It needs this in order to be able to optimize the selections and predictions that it makes as it tries to present its user with appropriate things to say as their next contribution to the conversation. Some of these items will require modification or editing before being output to the conversation partner, and so efficient online editing and prediction facilities are needed. It is therefore desirable for the AAC system to contain a model of conversation that gives the system a structure to operate within, a structure that indicates what type of step should be taken next, and what type of item is required next in the conversation. This will guide the system in its search for stored items and dynamic configuration of the system for modification or replacement of such items under the control of the user. The efficiency of the process and hence the responsiveness of the user in a live conversation situation is dependent on the ability of such a model to accurately guide the AAC system. Increasing emphasis will be placed on this modelling as AAC systems acquire larger stores of utterances, narratives and documents, including multimedia material. The need to manage and marshal this store of material for best use during conversation will mean that the conversation model within the AAC system will be called upon to make ever more accurate and relevant use of these resources. This paper discusses issues involved in modelling conversation within AAC in the context of real-time message-based communication (where messages can consist of utterances, texts, multimedia items or combinations of these) and proposes an integrated model to represent and underpin the process within an AAC system. It commences with an overview of methods that have been investigated for supporting effective conversation in AAC.

2. Supporting conversational interaction in AAC

A range of research projects has sought methods to support conversational interaction for people who use AAC. Utterance-based communication is discussed in Section 2.1, followed in Section 2.2 by methods for making it easier for people to enter or select message elements for constructing utterances in AAC systems. Section 2.3 discusses approaches that have been investigated for the storage and management of communication material for use in AAC, including topic-related material, humour, narrative and multimedia content. This leads into how topic progression within conversation in AAC has been considered (Section 2.4) and its subsequent appearance in AAC system design. AAC researchers have investigated how techniques developed within Natural Language Processing (NLP) and Generation (NLG) might be applied within AAC systems; NLP and NLG can assist in the production of well-formed utterances, for example, and this is discussed in Section 2.5.

Transactional communication, which is communication with the intention to accomplish a specific task, depends on the rapid and effective production of relevant utterances during live interaction; related AAC methods based on scripts, frames and schemata are hence described (Section 2.6). Successful transactional communication requires interaction unimpaired by delays or inappropriate conversational moves or violations of conversational maxims; research on maxim violation in AAC is therefore considered at this point (Section 2.7). Context awareness is also important in AAC and relevant to transactional communication; there is growing interest in VOCAs being able to offer support that is relevant to their current context, position or situation (Section 2.8). Further, in the design and operation of an AAC system, there are degrees of freedom in terms of how best to exploit its power and capacity and that of its user interface. There is also scope for flexibility and variety of expression, for example multiple ways of saying something, or different moods in which to say it. Consideration of these factors, in Sections 2.9 and 2.10, concludes this part of the paper and forms a prelude to the discussion on modelling of conversation for AAC that comes thereafter.

2.1 Utterance-based communication

Language can be seen as a means to achieve interactional goals (Goffman, 1959, 1981; Searle, 1969). It has also long been apparent among researchers and clinicians that AAC systems should be seen as means towards better communication and social interaction for people who use AAC (Kraat, 1985; Light, 1988). There is thus a strong motivation to improve the means by which people who use AAC can achieve interactional goals within conversation and social interaction. The effective use of whole utterances in a rapid and responsive manner within live conversation would be of clear benefit here.

Utterance-based communication in AAC requires the rapid retrieval from storage of whole phrases (utterances) for use in conversation. The structuring of the phrase store and its contents is an important factor in enabling rapid access to phrases when they are needed; this has proved to be a significant challenge in the development of utterance-based communication (Todman et al., 2008). For example, the Minspeak approach was originally designed with the aim of retrieval and use of whole utterances (Baker, 1982), although the focus later shifted towards word retrieval. Ever since, research on utterance-based communication has repeatedly shown the difficulty of achieving an effective solution for its use in AAC (Todman et al., 2008). It is recognized, however, that much of human speech is formulaic (Cheepen, 1988; Wray, 2001). AAC research has continued to pursue ways of exploiting this, the aim being to make it easier for utterances to be retrieved or generated automatically during conversation (Dempster et al., 2010a, 2010b; Higginbotham et al., 2002; Todman, 1999; Todman et al., 1999, 2008; Todman and Alm, 2003).

2.2 Speech acts: openers, closers and backchannels

Researchers in the early days of AAC were aware of the large difference in communication rate between naturally speaking people and people who were using AAC to interact (Foulds, 1980; Vanderheiden, 1983). This rate difference was seen as a major problem and impediment to effective conversation. Early research on word prediction, for example, was much motivated by this factor (Swiffin et al., 1985, 1987b). In the mid-1980s attention turned to trying to produce complete speech acts with individual selection actions (e.g. keystrokes) in order to facilitate rapid response and participation in conversational situations (Alm et al., 1987). A speech act was seen in this context as an utterance with a particular purpose. Conversation analysis was applied where possible to provide insights and directions for structuring conversation such that speech acts could be anticipated by the system and provided for use during interaction; the ultimate aim was to provide the right utterance at the right time (Alm et al., 1989a, 1992). At the outset it was evident that there were limitations to what could be achieved at that time with whole utterances, but some promising opportunities were beckoning.

Opening and closing stages of conversations are particularly formulaic and relatively predictable parts of an interaction (Krivonos and Knapp, 1975; Laver, 1981; Schegloff and Sacks, 1973). The CHAT (Conversation Helped by Automatic Talk) system was designed to contain recognizable and formulaic stages of conversation such as Greetings, Responses to Greetings, Smalltalk, Wrap-up Remarks and Farewells, hence accommodating the opening and closing phases of conversation (Alm et al., 1987, 1989a). Utterances were programmed into CHAT to fulfill these conversational stages, and sufficient numbers of phrases were included to allow the system to randomize its selection of appropriate utterances within individual stages (in order to avoid awkward repetition). It was also possible to set the system to choose utterances in a particular mood.

Various conversational strategies have been investigated with the aim of giving more possibilities and alternatives for the system to offer during conversation. Backchannel (feedback) communication is important in maintaining and controlling the flow of conversation, and in giving the speaker an indication of how the listener is feeling about the conversation (Yngve, 1970). Back-channels are short feedback words, phrases or sounds. They are normally delivered overlapping the utterances of the other speaker and, rather than creating an interruption, they assist the other speaker to carry on by conveying reactions to what they are saying. CHAT contained backchannel remarks (e.g. “Uh-huh”, “Okay”, “Great!”, “That’s too bad”), and a rapid-output feature that allowed general-purpose backchannel utterances to be quickly output. Generic utterances (Ball et al., 1999) and formulaic utterances (Wray, 2001) such as these gave the system a great deal of flexibility and variability for use in conversation. Trials of CHAT (using only opening/closing and feedback utterances) showed average speaking rates of 54 words/minute (Alm et al., 1992). Conversation was restricted to greeting, responding, giving feedback, and parting. This does not represent all of general conversation, of course, however participants judged the experience as more pleasant, natural and easy than attempting the same using a regular AAC system.

2.3 Further work on speech acts: topic, humour and narrative

Further experimentation on speech acts investigated the storage of specific topic-related statements in a prototype AAC system based on a text database (Arnott et al., 1988; Alm et al., 1989b). Stored utterances were categorized according to topic (e.g. family, sport) and type of speech act (e.g. inform, joke). These features were included in the TalksBac system, for people with aphasia, which allowed further categorization of statements on the basis of which conversation partner(s) they were most suited for (Waller et al., 1998). Humour was incorporated in the STANDUP system with the aim of helping people using AAC to use jokes and humour during interaction (O'Mara and Waller, 2003). Evaluation of STANDUP (Waller et al., 2009) with children with cerebral palsy showed that participants were able to use it to successfully generate novel puns, and work has continued in this area (Binsted et al, 2007; Black et al., 2007). Story-telling in AAC was also an important area of investigation (Waller et al., 1998, 2001; Waller, 2006; Black et al., 2010). Story-telling tools have been developed including ways to introduce a story, tell it at an appropriate pace and respond with feedback to listeners' comments. Being able to tell stories and recount narrative about one's past life is important for projecting one's identity and personality (Nelson, 1991; Polkinghorne, 1991). The use of multimedia broadens the scope for narrative and floor-holding available to the non-speaking person, helping them to engage a conversation partner in attending to audio-visual material as part of a conversation. Photographs can be used in story-telling (Balabanović et al., 2000) and video recordings can provide a means for illustrating personal histories and narratives (Hine and Arnott, 2002; Hine et al., 2003).

2.4 Topic progression in conversation

Topic in conversation shifts typically in a stepwise fashion, and topic-management in AAC needs to take this into account in order to help someone using AAC to negotiate such topic changes. A number of methods have been investigated with a view to facilitating movement between topics during utterance-based communication, such as semantic labelling of utterances, the use of hypertext and hypermedia, fuzzy information retrieval, and intersecting conversational perspectives.

A semantic network is a network which represents semantic relations between concepts. This is often used as a form of knowledge representation in artificial intelligence investigations. A prototype AAC system incorporating semantic networks was proposed with sets of sentences, phrases and stories input in advance by the user along with semantic labelling, such that the system would subsequently retrieve the pre-stored material in semantically appropriate contexts (Waller et al., 1990).

Hyperlinks in hypertext can be used to link different texts and it has been proposed that utterances stored in hypertext format could enable stepwise movement between topics (Alm et al., 1990). The creation of hypertext appropriate for individual users of AAC would be time-consuming and costly, however, and represents an obstacle to deployment. A hypermedia communication support system containing generic media material has, however, been seen to assist a person with dementia to communicate with a relative or carer. The hypermedia structure allowed flexibility and flow to occur in conversation despite the media material not being specific to any one user (Alm et al., 2007).

Fuzzy information retrieval allows for flexibility in storage and retrieval of information such that some approximation can be achieved in topic matching and categorization. In the context of utterance-based communication, this can make it easier to retrieve items from storage that have approximate rather than exact links to each other, so allowing movement to other topics that can be more or less closely related to each other. Experimentation in this area showed that conversations held using a fuzzy retrieval method were considered to be coherent to an acceptable degree (Alm et al., 1993).

In the TALK system, gradual perspective shifts could be used to move between topics, e.g. discussing different aspects of a topic that relate to past, present and future and using these different aspects to bridge to different topics (Todman et al., 1994). TALK evaluated positively in an AAC context, with a positive effect on perceived quality of conversation and rates of up to 64 words per minute being achieved (Todman et al., 1995, 1999, 2008). Higher word rates, of about 80 words per minute, were attained when greater use was made of extended personal narrative (Todman, 1999).

2.5 Natural Language Processing and Generation

2.5.1 Natural Language Processing (NLP)

Research on Natural Language Processing (NLP) investigates ways in which computers can analyse and manipulate language. Techniques developed in NLP have relevance to AAC in that they potentially yield means by which computers (and hence AAC systems) can automatically process language and improve the efficiency and efficacy with which parts of the communication process can be automated within AAC systems. A number of NLP techniques have been investigated in AAC (Langer and Hickey, 1998; Newell et al., 1998). The Compansion system was designed to process and expand spontaneous language constructions such that a group of uninflected content words could be converted automatically into a fully formed phrase or sentence (Demasco and McCoy, 1992; McCoy et al., 1998), for example, and syntax processing was investigated for use within predictive AAC interfaces (Swiffin et al., 1987a; Morris et al., 1991; Yang et al., 1990).

Work on the application of lexicon-based text retrieval techniques to AAC used a semantic lexicon derived from the WordNet database in the development of a prototype AAC system (WordKeys) based on message retrieval by keywords related to the semantic content of the message (Langer and Hickey, 1998). User evaluation of this approach showed that full text retrieval was an effective approach to message access in an AAC system. Research on NLP for AAC continues and recent work has investigated, for example, using speech recognition and a parser to display noun phrases spoken by a conversation partner to a person who uses AAC (Wisenburn and Higginbotham, 2008, 2009). This work indicated that the approach could be beneficial and may provide context-specific information that could improve rate and relevancy for the person using AAC, although it is dependent on the accuracy of speech recognition. In other work, context-based priming of word prediction (Higginbotham et al., 2009) was found to give marginally significant improvement in keystroke saving, although this did not necessarily translate into better overall task performance.

2.5.2 Natural Language Generation (NLG)

Natural Language Generation (NLG) is a technique which transforms data into natural language, usually for the purpose of automating the process of turning a set of facts and figures into readable text for a variety of audiences. Typically the input to an NLG system is raw data, and the output is well-formed text. NLG technology has been applied to AAC, to assist in story-telling and in participation in social dialogue (Reiter et al., 2009). The story-telling project uses information about a child's activities in school as its input. The information is logged automatically as the child moves through the school for classes and activities. This involves mobile (cellular) technology to track interactions with people and objects and locations during the school day. In addition, photographs and voice messages can be recorded for adding to the narrative which is generated. The output of the system is an account of the child's activities, which the child can then edit for use later in telling their family members about their day (Black et al., 2010; Black et al., 2011). This is a means of supporting personal narrative in AAC.

NLG is also being investigated with a view to assisting in the generation of conversational utterances (Dempster et al., 2010a). The user would enter a small amount of data and have that converted into extended, well-formed conversational contributions, thus assisting in the rapid generation of conversational utterances. The input for this system is data that the user records about their activities immediately after the activity occurs, in the form of single-word or short-phrase entries in a database. The data (typically answering the questions Where, When, Who with and What) are entered in preparation for speaking about it later. The data are then transformed by the system into a set of conversational utterances which the user can select from during subsequent conversation. Researchers in this area believe that rapid access to well-formed, contextually generated material created by NLG could lead to significant benefits for AAC users and their conversation partners (Dempster et al., 2010b).

2.6 Transactional communication

Transactional interactions, which are interactions aimed at achieving particular tasks such as placing an order for a meal in a restaurant or consulting a physician, can be seen as relatively formulaic and therefore somewhat predictable. Formulaic or predictable features in communication can be exploited in AAC; in this case a script/schema approach has been investigated as a way of applying anticipated sequences and structures to transactional interactions such that the system would retrieve and make available appropriate and relevant utterances for each stage of a particular interaction (Alm et al., 1995; Vanderheyden, 1995; Vanderheyden et al., 1996; Carpenter et al., 1997; Harper et al., 1998). In AAC, personalized scripts might be created such that persons who use AAC could use them to achieve particular tasks, such as shopping, visiting a physician or ordering food. A prototype AAC script system with six scripts was effective in trials, however it became clear that a larger repertoire of scripts was required for more general application of the concept (Dye et al., 1998).

Frames have also been applied in this context (Higginbotham and Wilkins, 1999; Higginbotham et al., 1999). A frame is a data structure that can be used to represent objects and situations; there are slots in the frame into which different items can be placed to represent the prevailing situation or communication context. Utterance templates can be stored in frames with slots being used for the insertion of different items from a database, thus giving a flexible tool for the generation of individual utterances for use in context during conversation. An AAC system, Frametalker, was constructed using the frame concept; word rates of more than 40 words per minute were achieved by participants during trials with Frametalker (Higginbotham et al., 2002). Features from TALK and Frametalker were combined in the Contact AAC system (File et al., 2003; Lunn et al., 2004), technology that was later commercialized for the AAC marketplace.

2.7 Violation of conversational maxims

Response time is important in conversation as is the avoidance of violations of conversational maxims such as lacking relevance, giving too much or too little information, or delaying response (Bedrosian et al., 2002, 2003; Hoag et al., 2004, 2008; McCoy et al., 2007, 2010a, 2010b). Experiments have been performed on conversational trade-off choices when an utterance-based system is used in goal-directed public situations. Quickly delivered messages with only partly relevant information were consistently found to be judged least favourably in experimental study, while quickly delivered messages with repetitive information were consistently most favourably received and met with much success in meeting their goals (McCoy et al., 2010a, 2010b). Other conversational trade-off choices (e.g. slow adequate message, fast inadequate message, fast excessive message) lay between the previous two. It was concluded that future utterance-based system design must provide ways to maximise the availability of situationally relevant pre-stored messages. Utterance-based technology must integrate seamlessly into AAC systems, and it must be possible to edit pre-stored messages easily for excessive or inadequate information, while supporting also the real-time construction of new messages. An AAC system should therefore be able to perform the necessary language processing, generation and prediction in real-time with appropriate (i.e. situationally relevant) utterances delivered promptly and correctly for use in conversation.

2.8 Context awareness in AAC

Context is important in AAC; a number of projects discussed thus far concern the issue of context or situationally relevant support for the person using AAC (e.g. Black et al., 2011; Higginbotham et al., 2009; McCoy et al., 2010a, 2010b; Wisenburn and Higginbotham, 2008, 2009; Waller et al., 1990). Some research has suggested using the Internet as a source of conversational material for use in AAC (Ashraf and Ricketts, 2003; Luo et al., 2007; Reiter et al., 2009). The idea is that the person using AAC would be able to say something like 'Have you heard about the election results?', where 'Have you heard about' would be a template and 'the election results' would be produced from information accessed on the Internet. The subsequent conversation would proceed with the person using AAC making use of a series of such templates with slots, with newly created slot-fillers derived from Internet news sites and similar sources. Challenges lie in how to harvest such material effectively and efficiently and how to present it to the person using AAC for easy inclusion in their conversation.

The iconCHAT system was developed as a prototype message construction VOCA based on semantic frames and graphical (icon) input selection (Patel et al., 2004, 2006; Patel, 2010). The aim was to give an alternative to syntax-based formulation of sentences, through the semantic organization of message construction. In further development, geographical sensing using GPS (global positioning system) was applied in an exploration of location-specific vocabulary usage that could inform the vocabulary content of iconCHAT (Patel and Radhakrishnan, 2007). Continuing advances in sensor technology are likely to make it easier to automatically identify physical and social contexts within which VOCAs are being used, so context awareness is likely to be an area of continuing and growing importance within AAC.

2.9 Degrees of freedom in design and operation of an AAC system

In the design and development of an AAC system, there are choices and compromises to be made concerning how to use the available computing power of the system and the interface capacity between the system and the person using it. A number of degrees of freedom exist in the design and operation of an AAC system, such as size of vocabulary and language units contained therein, the practical capacity of the selection space(s) available to the user, the selection methods and the selection units available. Vocabularies and system memories can contain characters, words, phrases, paragraphs or whole narratives or texts, including multimedia material (e.g., photographs, recordings) that can be used during conversation. Similarly the selection units used can be symbols, characters, words, phrases, paragraphs, documents or images representing these things. A wide range of selection methods (e.g. scanning, coding, direct selection) and interface technologies for input and output exist (Beukelman and Mirenda, 2005; Higginbotham et al., 2007b).

Given that design choices of this nature will need to be made, it is necessary to consider how to balance the number of things that can be said and in how many ways they can be said. An AAC system could be designed to be able to say many things in a few ways, or a few things in many ways, or a few things in a few ways. Saying many things in many ways is more challenging and will be constrained by prevailing technical limitations, selection space and interface technology. Increasing the number of things that can be said and the ways of saying them increases the demands on the system and the number of selection choices that the person using it might have to make, which is undesirable. If the system could automatically and accurately narrow down selection choices, the cognitive and physical load on the person using the system would be reduced and conversation made easier for all those involved in it. This is difficult to achieve, however; the limitations of computer-based language processing and generation are a principal constraint here.

2.10 Flexibility and variety of expression in AAC

An early example of having different ways to say things was manifest in the CHAT (Conversation Helped by Automatic Talk) project, where a variety of utterances were available to say essentially the same thing (Alm, 1988; Alm et al., 1987, 1989a). CHAT could make a random choice each time from a list of suitable alternative utterances for the next expected conversational move. This randomisation was done to avoid repetitions of identical utterances in any category. The first Greeting to come out might therefore be “Hi, how are you?”, while the Greeting produced in the next conversation might be “Hello there. How’s it going?”. CHAT also supported a range of moods (e.g. Polite, Informal, Humorous, Angry); a user could select an appropriate mood whereupon subsequent utterances were worded according to that mood. Subsequent work on emotion in synthetic speech investigated how affect could be simulated in speech output from a VOCA so that mood expressed in an utterance could be emphasised by features of the synthetic voice (Murray et al., 1991; Murray and Arnott, 1995, 2008). This adds to the variability and flexibility of the output in that synthesis-by-rule can change how the output sounds (affect). An AAC system should be able to accommodate variety and flexibility in its content and the ways in which the content can be used. It will need to be able to reflect the emotional state (mood) of the user, for example, and move through different emotional states, while controlling speech synthesisers to simulate affect in VOCA speech output.

3. Modelling of conversation for AAC

3.1 Need for modelling of conversation

The aim of utterance-based AAC, as described above, is to try to enable a person who uses AAC to produce appropriate utterances at the right time for use in conversation. The prototype systems above have demonstrated that significant advances can be made in responsiveness and communication rate if appropriate utterances can be offered in a timely manner by the AAC system.

There is a compromise that has to be achieved here. There is an obvious difficulty with trying to predict accurately and precisely, during a conversation, the next statement that a person using AAC needs to make. Life is often unpredictable, and machine understanding is not sufficiently developed for a computer to comprehend discourse, so it cannot make all the higher-level decisions and responses that are required during a live conversation. However it is necessary for an AAC system to provide as much support as possible to its user during conversation. For this, a model of conversation is needed so that the system can have a computer-based structure of what is happening during the conversation, thus giving a "scaffold" or landscape within which the AAC system can make a best attempt to predict appropriate conversation acts, moves and utterances for a non-speaking person to use during live conversation. This should be achieved in real-time without delays that might impede the progress of the conversation.

3.2 Previous work on modelling of conversation for AAC

A number of attempts have been made to create a model of conversation that would be helpful in designing AAC systems. These efforts took as a starting point the work of the conversation analysts (Garfinkel, 1967; Sacks et al., 1974; Levinson, 1983; Atkinson and Heritage, 1984; Tannen, 1984; Cheepen, 1988; Hutchby and Wooffitt, 1988; Sacks, 1995; Lerner, 2004; Schegloff, 2007; Heritage and Clayman, 2010), the systemic functional linguists (Halliday, 1985; Halliday and Martin, 1981) and the synthesizing of much of this work by Clark (Clark, 1996). None of this basic research was done with AAC specifically in mind, but it has provided a foundation on which to build models of conversation particularly suited to assisting our thinking about improving AAC systems. Conversation models that were produced by Alm (1988) and Alm et al. (1987), Todman and Alm (2003) and Hedvall and Rydeman (2010a, 2010b) have illuminated different aspects of AAC conversation.

The model of Alm et al. (Figure 1) emphasized the overall contour of the conversation and the stages through which conversation moves. It was derived from the work of the early conversation analysts, who demonstrated that conversation, far from being incoherent and incapable of careful analysis, has a definite structure. The model shows that there are normally two opening and two closing phases to an encounter. A potentially useful feature of these phases for the AAC user is that they tend to be quite formulaic.

Todman and Alm's model (Figure 2) concerned itself with important pragmatic aspects of conversational interaction and social impact. The model portrayed the structure of an AAC system and the social impact achievable with it. It showed that any AAC system with pre-stored text would always need a facility for creating unique utterances, which although slower, would allow the user to accomplish important pragmatic tasks. It also made explicit a set of important interactional goals that conversation should accomplish, whether carried out with an assistive system or not.

Hedvall's "Activity Diamond" (Figure 3) foregrounded the role of the speaker's physical and social environment (Hedvall and Rydeman, 2010b). A dyadic use of the Activity Diamond is seen in Figure 3, in which both conversants (the person using AAC and their personal assistant) and their environments (contexts) are shown during an activity (shopping). The model "portrays a human activity system" that incorporates a subject (e.g. the person using AAC), an object (e.g. a shopping activity, in which the AAC user is shopping), an artefactual/natural environment (material and immaterial artefacts and nature around the subject), and a human environment (people and society around the subject). This example is shown in the Activity Diamond in the left-hand half of Figure 3.

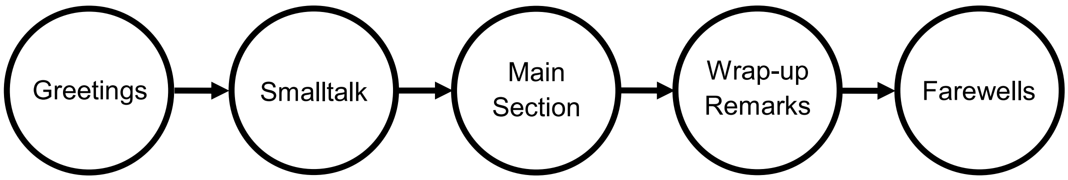

Figure 1. *Conversation contour. The overall contour of a conversation shows the stages through which conversation moves, starting with Greetings and concluding with Farewells. (After Alm, 1988)*

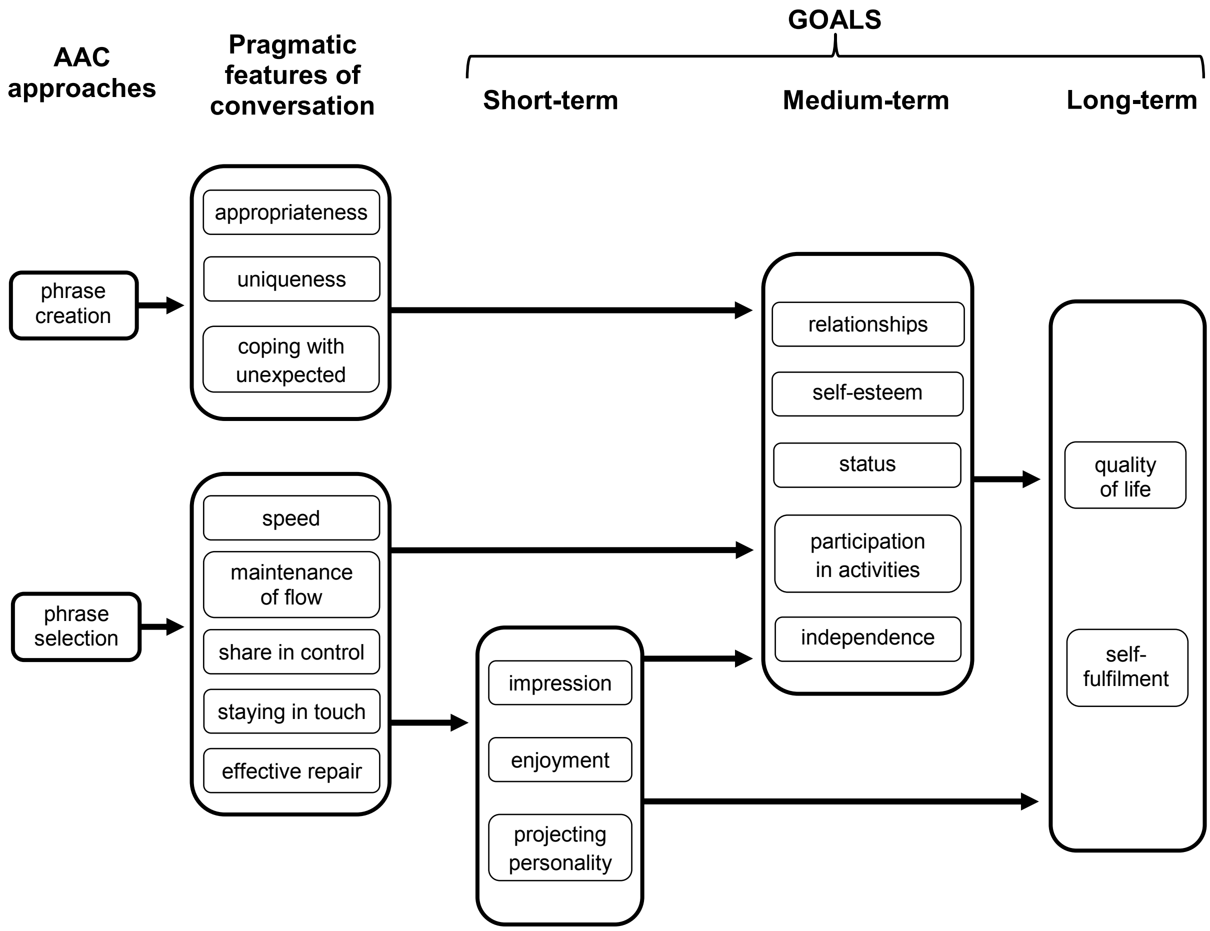

Figure 2. *Social impact model. A model of conversation emphasizing social impact, which links pragmatic features of conversation with user goals and AAC approaches. (After Todman & Alm, 2003)*

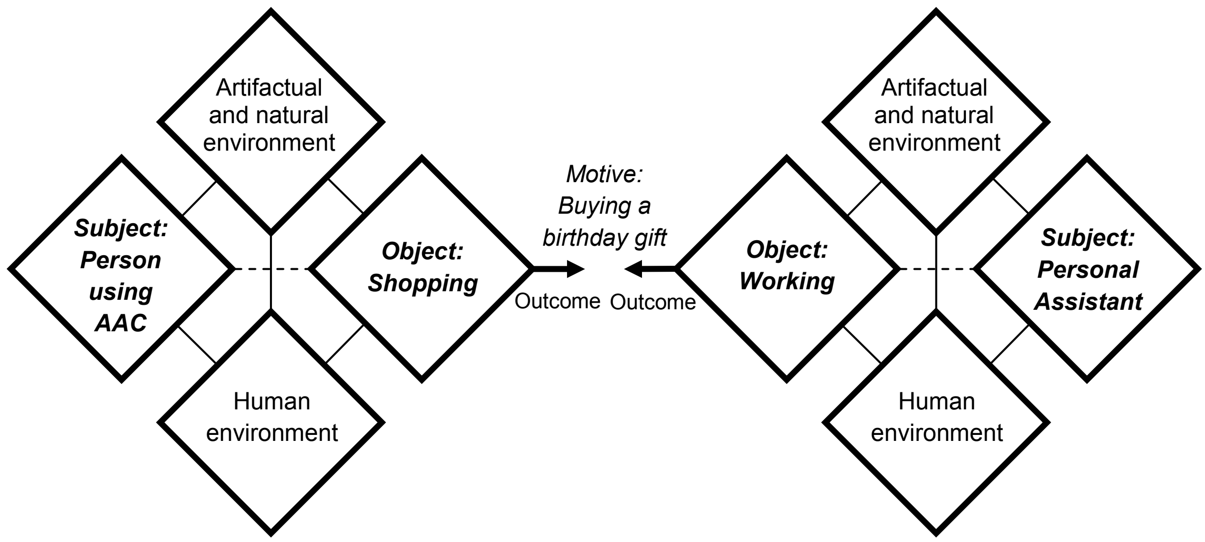

Figure 3. *The Activity Diamond. A dyadic use of Hedvall's Activity Diamond is shown here that represents a person using AAC with their personal assistant during a shopping expedition. (After Hedvall & Rydeman, 2010b)*

An equivalent Activity Diamond for the personal assistant of the person using AAC is shown in the right-hand half of the dyad in Figure 3, where the personal assistant's object is shown as 'working'. The interaction between the person using AAC and the assistant is thus portrayed, including the contexts that they are in and the objects of their activity.

We would argue that none of these are complete and competing models, but rather each represents a different way of looking at a conversation. As such each provides useful insights. The model in Figure 1 suggested that significant parts of conversation may be stereotypical and thus relatively easy to provide quickly to AAC users. The model in Figure 2 made explicit the pragmatic, interactional aims of conversation, only a small part of which has to do with the transfer of objective information. The model in Figure 3 points out that the context of an interaction is important, both the speakers' physical surroundings and the social context in which they meet.

The differences between the models can be illustrated by showing the ways in which each one could contribute to the design of AAC systems. The model in Figure 1 makes it clear that having quickly available pre-packaged utterances to start off and finish an interaction would be very useful. The model in Figure 2 shows why this is important, helping the user to achieve short-term, medium-term and long-term goals in how they present themselves. The model in Figure 3 suggests that, if technology could provide the user with automatically captured information about their physical and social environment, this could be very helpful in a conversation.

The first two of these models represent just one side of the conversation. This is to be expected since their purpose is to allow us to examine the potential output of an AAC system being used by one of the conversation participants. The model in Figure 3 goes further in allowing representation of both sides of the exchange. The version shown here does this, as it shows two activity diamonds in a dyadic formation, one for each participant in the exchange.

3.3 A new model of conversation for AAC

One problem with these models is that they do not separate out in any detail the generic aspects of a conversation and features that are specific to AAC. Having a separate reference to AAC in a model would be useful in bridging the gap between theory and practice for communication system designers. We have taken the model in Figure 1 as a starting point, since it attempts the most explicit coverage of the structure of conversation, and have produced a new model of conversation which we would argue is at the same time more comprehensive at the top level, and more detailed about AAC at the bottom level. Figure 4 sets out the top-level view of the model, which shows how a conversation is structured, from the point of view of one of the conversants. Figure 5 is the AAC-specific part of the model, and represents a 'zooming in' on any of the individual stages from Figure 4, showing tools that are available to an AAC user to accomplish the particular conversational move for that stage.

At the highest level, a conversation consists of an opening phase, a central part of the interaction, and a closing phase. Deliberately leaving out the opening phase is a rule violation that can be employed to convey urgency or anger (e.g. A walks up to B with no preliminaries and says "What did you tell Bill about me?"). Leaving out the closing phase would in almost any circumstance be perceived as bizarre, but again could be used to convey a strong sense of urgency or a negative emotion (e.g. "What did I say? She just turned on her heels and walked off!"). Grice was first to try to systematize the way that people will deliberately violate conversational rules as a way of conveying meaning (Grice, 1975). It is possible to conceive of an interaction which consists only of opening and closing moves ("Hi, must rush - how's it going? OK, see you later."), but as with this example an apology for the truncated interaction might be expected.

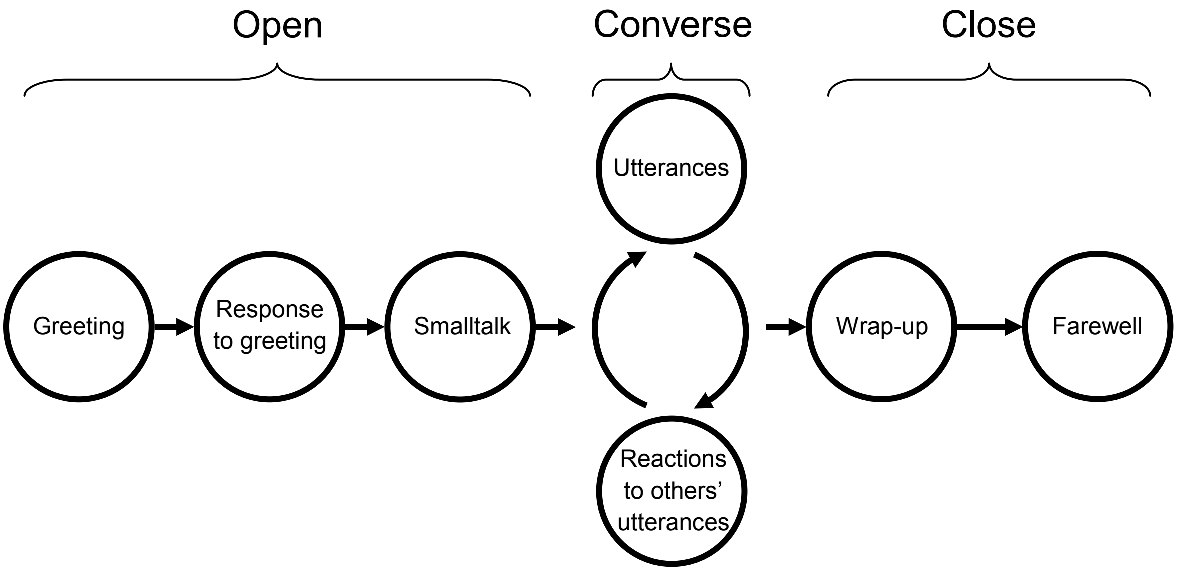

Figure 4. *Conversation stages. There are three opening stages and two closing stages surrounding a central part where the main discussion takes place.*

Figure 4 shows three opening stages (Krivonos and Knapp, 1975; Laver, 1981; Rosner, 1981; Sacks, 1995) and two closing stages (Schegloff and Sacks, 1973; Knapp et al., 1973; Laver, 1981; Sacks, 1995) of an encounter. The central part of an interaction is, of course, widely diverse, but what can be said is that it will consist either of utterances or reactions to the utterances of others. In a sense, after the initial utterance of an interaction, all the other utterances tend to be reactions to others' utterances. This is what makes up the coherence of a conversation, where each utterance, from any party, follows on from a previous utterance. An exception is when an external event, or an interruption from someone who joins the conversation, occurs that is then commented on by the conversants (which is part of the point made by Hedvall about context playing a role in a conversation (Hedvall and Rydeman, 2010a, 2010b)). Frequently the reaction to others' utterances consists of a backchannel remark, which is intended not as an interruption but as feedback.

The diagram shown in Figure 5 represents the process of producing an item of conversational output with an AAC system. At all of the individual conversational stages, from Greeting to Farewell, shown in Figure 4, the set of tools represented in Figure 5 can be used to generate an item for output.

There are two principal sub-processes in Figure 5, representing Creation and Selection. The Creation process contains tools for composing output messages element-by-element, for example assembling words or phrases character-by-character using assistive input tools such as coded input, scanning selection or prediction. Predictive routines have been developed that optionally may aid the process, predicting forthcoming characters or words, disambiguating between selection options on the basis of items already entered, or predicting likely whole utterances for the user to select from. The Selection process contains tools for selecting appropriate items from a Repository. The Repository is a store of communication material such as utterances, images, video and audio recordings, text and multimedia documents. The Selection tools can include the use of scaffolding to guide the extraction of material from the Repository, so that a script, for example, could be used to guide the selection of other items from the Repository appropriate for use in the current stage of conversation. Various 'scaffolding' schemes have been proposed thus far, normally based on predicting likely utterances given the current one, as detailed in Section 2 above.

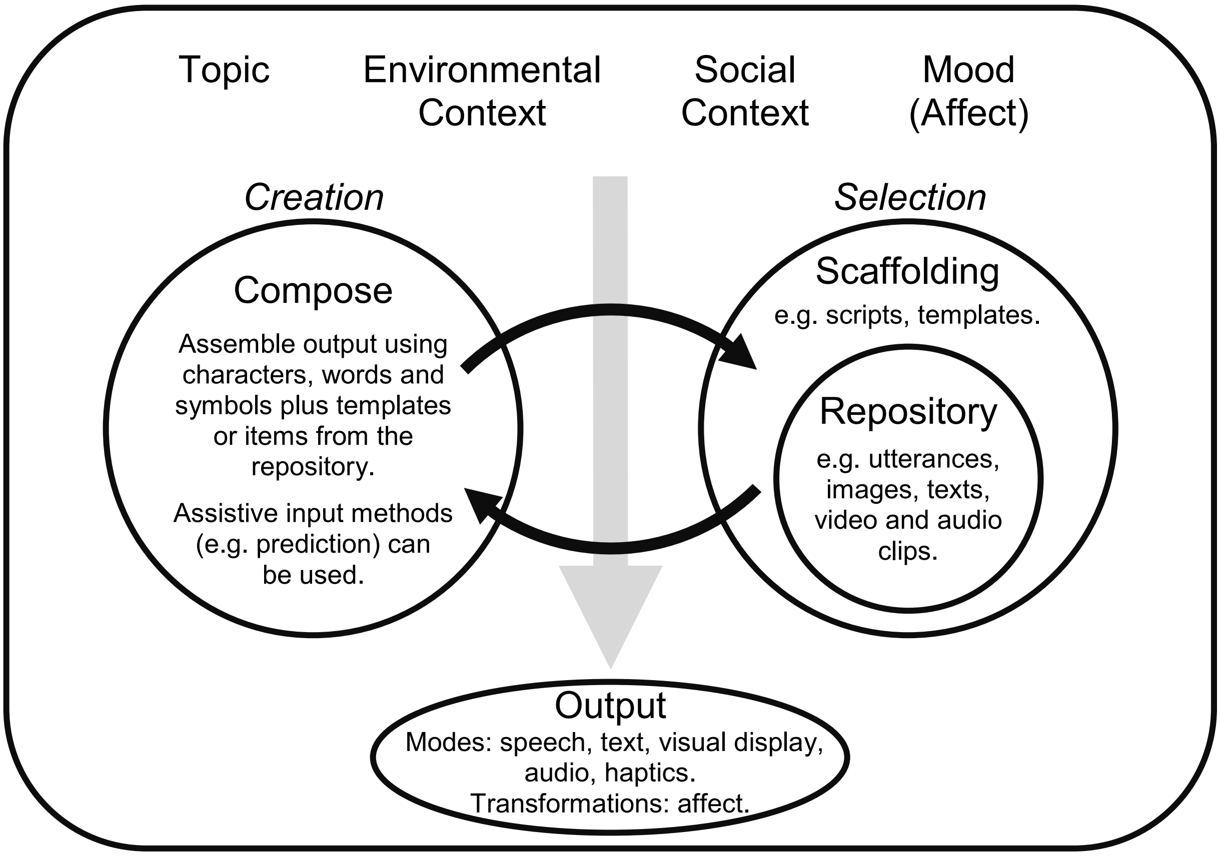

Figure 5. *AAC communication tools. These are communication tools which can be used in the conversation stages shown in Figure 4 to support the production of output (e.g. utterances) for use in the current stage of conversation. Prevailing conversational conditions (topic/context/mood) are shown at the top of the rectangle and these pervade and inform the other elements in the toolbox so that appropriate output is produced for these conditions. The central vertical arrow represents progress through the process of producing an item for output, culminating in the output of that item using appropriate output mode(s); the production process can cycle between Creation and Selection as the output item is being produced.*

The process of producing a conversational output item can cycle between the Creation and Selection processes, with, for example, Selection producing a template phrase then Compose tools being used to enter words into empty slots in that template. The central vertical arrow in Figure 5 represents the progress of the person using AAC through the process of producing an item for output, culminating in the output of that item using appropriate output mode(s). This production process may be short and rapid, if the item is brief (e.g. a wrap-up remark), or long and involved (e.g. if scaffolding is required to construct an item such as a narrative and/or several cycles through the creation-selection cycle are performed in its construction). For instance, if the user wished to say ‘Uh huh’, this would come from the Selection mode, and might be picked out from an always available set of feedback remarks. This might take one selection action to accomplish. If, on the other hand, the user wanted to add the word ‘absolutely’ here, they might have to compose that word letter-by-letter, assisted by a prediction tool. This could take, say, four selections to accomplish: three selection actions to input the individual letters ‘a’, ‘b’, ‘s’, followed by one action to select the word ‘absolutely’ from a prediction menu. Putting the whole phrase together would thus be a combination of rapid and slower composition phases. The model thus accommodates the production and output of items of different lengths and complexity. It can also accommodate different types of output, such as multimedia items like video clips, and the mixing of media such as template phrases containing words as well as slots for symbols or images.

Prevailing context, topic and mood are represented in Figure 5, providing a pervading back-drop to the processes. Physical context embraces the physical location, position or environment of the person who is using the AAC system while social context represents the human or social situation that the person is in, including information about whom they are talking to at the time. The topic is the current topic of conversation, which informs the processes of Creation and Selection by promoting items relevant to that topic. The mood is the prevailing emotion or mood that the person using AAC is wishing to express or impart in the conversation. Mood could influence the Creation and Selection processes by adjusting the availability of items relevant to that mood.

In many cases the output mode will be synthetic speech, but other modes such as visual display, haptic interface and audio (e.g. for multimedia sound) are also possible. There may optionally be transformation of the output occurring at this point, such as modification of synthetic speech to evoke affect, thus reflecting the current mood.

3.4 Summary

The model described here (Figures 4 and 5) shows the stages through which a person using AAC will move during a conversation. He or she will be seeking to participate in conversation with a conversation partner and will step from stage-to-stage through the conversation using the AAC system to interact with that partner. The model gives the AAC system a structure that guides it regarding the type of output item that will be needed in the upcoming stage of the conversation; this helps the system to focus down on items that are appropriate for that stage of the conversation so that it can prepare to offer these to the user for selection and use. A model of this nature could also assist those involved in AAC research and development to envisage and analyze the process that an AAC system must go through as conversation unfolds.

4. Possible future developments

4.1 Context awareness

Context awareness, an aspect of conversation modelling which has been discussed, has the potential to develop significantly due to technical advances. Geographical position sensing can be used (Patel and Radhakrishnan, 2007) to indicate geographical context and other methods of physical context sensing are also feasible. Tiny low cost sensors can create an environment which is continuously monitoring itself. RFID (radio frequency identification) tags are small devices which can be attached to objects to broadcast information about themselves to anyone nearby. Sensor networks and pervasive technology can be used to detect a person's presence and movement in spaces (Gil et al., 2007) and their interactions with objects. Such technology is moving us towards what has been called 'the internet of things', potentially a global network of physical objects as pervasive and ubiquitous as the worldwide web itself. Work so far on context awareness in AAC has noted the importance of this in everyday conversation, and indicated that enabling a VOCA to respond to its physical or social environment would be advantageous. Technical means for context awareness are continuing to emerge and engaging the attention of AAC researchers (Black et al., 2010). This is likely to be an area of increasing interest and importance within AAC.

4.2 Co-construction of conversation

It has been acknowledged in the discussion of possible models of conversation for AAC that it is only the one side of the conversation which is normally accounted for in the detailed models (the very high-level model shown in Figure 3 does cover two people interacting). This is no doubt because modelling one side of a conversation is challenging enough, given its variable nature. Attempting to include both conversants increases the difficulty. However, it is known that conversation is in fact always a 'co-constructed' activity (Clark, 1996). Involving the input from the other speaker is an idea which would seem to have promise, but how to bring about the practical implementation of it can be less clear. Where communication is over a computer network, or good quality speech recognition is possible, however, the other person's contributions are available for analysis and use in the construction of other messages.

Co-construction of messages has been investigated in work on collaborative writing in augmentative devices (McKinlay et al., 1993), and more recently in experiments with conversation partners providing guesses about intended messages which would be included in the word prediction interface for a person using AAC (Roark et al., 2011). By providing word completions and predictions to a person using AAC, participants (conversation partners) in this research were able to contribute to reducing the numbers of keystrokes required in text entry. These human-assisted keystroke reductions did not match the reductions achieved by computer-based n-gram language models, however. Higginbotham et al. (2007a) studied co-construction, with communication partners in experimental conditions actively participating in the construction of the augmented communicator's output. Work has also been done on using speech recognition to capture the other speaker's communication (Wisenburn and Higginbotham, 2008, 2009) (see also Section 2.5).

An interesting possibility is to have the AAC system being physically used by both the non-speaking person and the conversation partner. This would truly be a co-constructed dialogue. A prototype system was developed for cross-language communication which took this approach and was, in fact, inspired by AAC systems (Iwabuchi et al., 2001). It may be possible to take this idea further, with the starting point being that interacting with language is a type of game (Wittgenstein, 1953). One research direction would be to examine to what extent a system which facilitated communication, and which was controlled equally by the non-speaking person and the communication partner, would be acceptable and capable of creating a satisfying interaction.

4.3 Emotion-based communication

It is clear that expressing emotions is an important part of everyday conversation, and that this can be difficult for someone who uses AAC. Research has been done on adding emotional effects to synthetic speech for AAC applications (Murray et al., 1991; Murray and Arnott, 1995, 2008), attempting an output transformation to add affect to text utterances. The widespread adoption of concatenative speech synthesis has constrained this process somewhat; adding voice variability to this type of speech is more difficult than with older formant-based synthesisers, although numerical methods to modify this type of speech have been studied (e.g. Hofer et al., 2005; Hamza et al., 2004). Research on methods for improving expressive speech synthesis continues (Schröder, 2009). Some commercial VOCAs now offer support for expression of mood and emphasis in the speech output that they produce (e.g. DynaVox® Xpress™, DynaVox® Tango!™).

A more radical approach to expressive speech has been made in the form of what is really a prototype designed to provoke discussion; in this project on interaction with synthetic speech the tone of voice is considered to be what matters, not the vocabulary (Cook and Pullin, 2010; Pullin and Cook, 2010). A very small vocabulary of words (e.g., "yes", "no", "hello" and "really"), chosen on the basis that the perceived meaning of each word depends greatly on how it is pronounced, is made available within an experimental control space similar to an AAC interface (e.g., joystick, small number of switches). Expressive range is dependent on skill in controlling the synthetic vocal inflection, via the interface, in order to infuse the words with meaning. This is a novel approach to expression that might become an optional feature for future VOCAs. In addition, although not presented as a full AAC prototype, this system does raise an interesting question: what if instead of modelling conversational structure we tried to devise ways to communicate effectively using intonation alone? How far could this take us?

5. Conclusion

This paper has shown a model of conversation as seen from within an AAC system; it models the production of one side of conversation, the side produced by the person using the AAC system. Modelling the process in this way gives a structure around which an AAC system can organize the support and guidance that it offers to the person using the system. Modelling of this nature can also promote analysis and understanding of what an AAC system is likely to be called upon to do in a conversational situation.

Much research has been seen in the field of AAC; a substantial amount of this is in the area of conversation, yet, apart from some notable exceptions described in this paper, there has been relatively little attention given to the actual modelling of the process of conversation. Given the importance of social and functional interaction for people who use AAC, it is important that more attention be given to this area in order to enhance understanding and encourage further improvement in conversation support for Augmentative and Alternative Communication.

Acknowledgements

The authors gratefully acknowledge the assistance of Dr. Iain R. Murray in reading and commenting on the manuscript during its preparation. The first author received support from RCUK Digital Economy Research Hub EP/G066019/1 “SiDE: Social Inclusion through the Digital Economy” during manuscript preparation.

References

Alm, N., 1988. Towards a conversation aid for severely physically disabled non-speaking people. Ph.D. Thesis. University of Dundee, Dundee.

Alm, N., Arnott J.L., 1998. Computer-assisted conversation for non-vocal people using pre-stored texts. IEEE Transactions on Systems, Man and Cybernetics: Part C 28 (3), 318-328.

Alm, N., Arnott J.L., Newell A.F., 1989a. Discourse analysis and pragmatics in the design of a conversation prosthesis. Journal of Medical Engineering and Technology 13 (1/2), 10-12.

Alm, N., Arnott J.L., Newell A.F., 1989b. Database design for storing and accessing personal conversational material. In: Proceedings of the RESNA 12th Annual Conference (RESNA ’89), New Orleans, Louisiana, USA, 25-30 June, pp.147-148.

Alm, N., Arnott J.L., Newell A.F., 1990. Hypertext as a host for an augmentative communication system. In: Proceedings of the European Conference on the Advancement of Rehabilitation Technology (ECART), Maastricht, The Netherlands, 5-8 November, pp.14.4a-14.4b.

Alm N., Arnott, J.L., Newell A.F., 1992. Prediction and conversational momentum in an augmentative communication system. Communications of the ACM 35 (5), 46-57.

Alm, N., Dye, R., Gowans, G., Campbell, J., Astell A., Ellis, M., 2007. A communication support system for older people with dementia. IEEE Computer 40 (5), 35-41.

Alm N., Morrison A., Arnott, J.L., 1995. A communication system based on scripts, plans, and goals for enabling non-speaking people to conduct telephone conversations. In: Proc. of the IEEE Conference on Systems, Man & Cybernetics, Vol.3, Vancouver, BC, Canada, 22-25 October, pp.2408-2412.

Alm, N., Newell, A.F., Arnott J.L., 1987. A communication aid which models conversational patterns. In: Proceedings of the 10th Annual Conference on Rehabilitation Technology (RESNA '87), San Jose, California, USA, 19-23 June, pp.127-129.

Alm, N., Nicol, M., Arnott, J.L., 1993. The application of fuzzy set theory to the storage and retrieval of conversational texts in an augmentative communication system. In: Proceedings of the RESNA '93 Annual Conference, Las Vegas, Nevada, USA, 12-17 June, pp.127-129.

Arnott J.L., Alm, N., Newell A.F., 1988. A text database as a communication prosthesis. In: Proceedings of the International Conference of the Association for the Advancement of Rehabilitation Technology (ICAART 88), Montréal, Canada, 25-30 June, pp.76-77.

Ashraf, S., Ricketts, I.W., 2003. Automated vocabulary collection to allow topical conversation for non-speaking people. In: Proceedings of HCI International (Universal Access in HCI: Inclusive Design in the Information Society), Crete, Greece, 22-27 June, Vol.4, pp.181-185.

Atkinson, J., Heritage, J. (Eds.), 1984. Structures of social action - studies in conversation analysis. Cambridge University Press, Cambridge.

Baker, B., 1982. Minspeak: A semantic compaction system that makes self expression easier for communicatively disabled individuals. Byte 7 (9), 186-202.

Balabanović, M., Chu, L.L., Wolff, G.J., 2000. Storytelling with digital photographs. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2000), The Hague, Amsterdam, Netherlands, 1-6 April, pp.564-571.

Ball, L.J., Marvin, C.A., Beukelman, D.R., Lasker, J., Rupp, D., 1999. Generic talk use by preschool children. Augmentative and Alternative Communication 15 (3), 145–155.

Bedrosian, J.L., Hoag, L.A., McCoy, K.F., 2003. Relevance and speed of message delivery trade-offs in Augmentative and Alternative Communication. Journal of Speech, Language, and Hearing Research 46 (4), 800-817.

Bedrosian, J.L., Hoag, L.A., McCoy, K.F., Higginbotham, J., 2002. Communication rate and the use of utterance-based technology. In: Proceedings of the 10th Biennial Conf. of the International Society for Augmentative and Alternative Communication (ISAAC 2002), Odense, 10-15 August, pp.129-130.

Beukelman, D., Mirenda, P., 2005. Augmentative and alternative communication: Supporting children and adults with complex communication needs. Edition 3. Paul H. Brookes Publishing Co., Baltimore, MD.

Binsted, K., Bergen, B., Coulson, S., Nijholt, A., Stock, O., Strapparava, C., Richie, G., Manurung, R., Pain, H., Waller, A., O'Mara, D., 2007. Computational humor. IEEE Intelligent Systems 21(2), 59-69.

Black, R., Reddington, J., Reiter, E., Tintarev, N., Waller, A., 2010. Using NLG and sensors to support personal narrative for children with complex communication needs. In: Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies, Los Angeles, California, USA, 5th June, pp.1-9.

Black, R., Waller, A., Ritchie, G., Pain, H., Manurung, R., 2007. Evaluation of joke-creation software with children with complex communication needs. Communication Matters 21 (1), 23-27.

Black, R., Waller, A., Tintarev, N., Reiter, E., Reddington, J., 2011. A mobile phone based personal narrative system. In: Proceedings of ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2011), Dundee, Scotland, UK, 24-26 October, pp.171-178.

Carpenter, T., McCoy, K., Pennington, C., 1997. Schema-based organization of re-usable text in AAC: user-interface considerations. In: Proceedings of the RESNA '97 Annual Conference, Pittsburgh, Pennsylvania, USA, 20-27 June, pp.57-59.

Cheepen, C., 1988. The predictability of informal conversation. Pinter Publishers, London.

Clark, H.H., 1996. Using language. Cambridge University Press, Cambridge.

Cook, A., Pullin, G., 2010. Insights from the six speaking chairs. In: Proceedings of the 14th Biennial Conference of the International Society for Augmentative & Alternative Communication (ISAAC 2010), Barcelona, Spain, 24-29 July, Paper O.168.

Demasco, P., McCoy, K., 1992. Generating text from compressed input: an intelligent interface for people with severe motor impairments**.** Communications of the ACM 35 (5), 68-78.

Dempster, M., Alm, N., Reiter, E., 2010a. Automatic generation of conversational utterances and narrative for Augmentative and Alternative Communication: A prototype system. In: Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies, Los Angeles, California, USA, 5th June, pp.10-18.

Dempster, M., Alm, N., Reiter, E., 2010b. The application of Natural Language Generation to conversational communication systems for non-speaking people. In: Proceedings of the 14th Biennial Conference of the International Society for Augmentative & Alternative Communication (ISAAC 2010), Barcelona, Spain, 24-29 July, Paper O.005.

Dye, R., Alm, N., Arnott, J.L., Harper, G., Morrison, A., 1998. A script-based AAC system for transactional interaction. Natural Language Engineering 4 (1), 57-71.

File, P., Todman, J., Higginbotham, J., Lesher, G.W., Moulton, B.J., Alm, N., 2003. Contact: A communication aid based on pre-prepared phrases. In: Assistive Technology - Shaping the Future, G. Craddock et al. (Eds), IOS Press, Amsterdam, pp.304-308.

Foulds, R., 1980. Communication rates for nonspeech expression as a function of manual tasks and linguistic constraints. In: Proceedings of the International Conference on Rehabilitation Engineering, Toronto, Canada, 16-20 June, pp.83-87.

Garfinkel, H., 1967. Studies in ethnomethodology. Prentice Hall, New Jersey.

Gil, N., Hine, N., Arnott, J.L., Hanson, J., Curry, R., Amaral, T., Osipovic, D., 2007. Data visualisation and data mining technology for supporting care for older people. In: Proceedings of the ACM SIGACCESS International Conference on Computers and Accessibility (ASSETS 2007), Tempe, Arizona, USA, 15-17 October, pp.139-146.

Goffman, E., 1959. The presentation of self in everyday life. The Penguin Press, London.

Goffman, E., 1981. Forms of talk. University of Pennsylvania Press, Philadelphia.

Grice, P., 1975. Logic and conversation. In P. Cole & J. Morgan (Eds.), Syntax and Semantics, Vol. 3. Academic Press, New York.

Halliday, M.A.K., 1985. An introduction to functional grammar. Edward Arnold, London.

Halliday, M.A.K., Martin, J., 1981. Readings in systemic linguistics. Batsford, London.

Hamza, W., Bakis, R., Eide, E.M., Picheny, M.A., Pitrelli, J.F., 2004. The IBM expressive speech synthesis system. In: Proceedings of Interspeech-2004-ICSLP, 8th International Conference on Spoken Language Processing, Korea, 4-8 October, pp.2577-2580.

Harper, G., Dye, R., Alm, N., Arnott, J.L., Murray, I.R., 1998. A script-based speech aid for nonspeaking people. Proceedings of the Institute of Acoustics 20 (6), 289-295.

Hedvall, P-O., Rydeman, B., 2010a. A cultural-historical design approach to augmentative and alternative communication. In: Proceedings of the 14th Biennial Conference of the International Society for Augmentative & Alternative Communication (ISAAC 2010), Barcelona, Spain, 24-29 July, Paper O.367.

Hedvall, P-O., Rydeman, B., 2010b. An activity systemic approach to augmentative and alternative communication. Augmentative and Alternative Communication 26 (4), December, 230-241.

Heritage, J., Clayman, S.E., 2010. Talk in action: interactions, identities and institutions. Wiley-Blackwell, Boston.

Higginbotham, D.J., Bisantz, A.M., Sunm, M., Adams, K., Yik, F., 2009. The effect of context priming and task type on augmentative communication performance. Augmentative and Alternative Communication 25 (1), 19-31.

Higginbotham, D.J., Kim, K-E., Scally, C., 2007a. The effect of the communication output method on augmented interaction. Augmentative and Alternative Communication 23 (2), 140-153.

Higginbotham, D.J., Lesher, G.W., Todman, J., File, P., Wilkins, D., 2002. Utterance based communication: theory, research and design. Short course presented at the Annual Conference of the Rehabilitation Engineering Society of North America (RESNA 25th International Conference), Minneapolis, Minnesota, USA, 29th June.

Higginbotham, D.J., Shane, H., Russell, S., Caves, K., 2007b. Access to AAC: present, past, and future. Augmentative and Alternative Communication 23 (3), 243–257.

Higginbotham, D.J., Wilkins, D., 1999. Frametalker: A system and method for utilizing communication frames in augmented communication technologies. US Patent no. 5,956,667. US Patent & Trademark Office.

Higginbotham, D.J., Wilkins, D.P., Lesher, G.W., 1999. Frametalker: a communication frame and utterance-based augmentative communication device. In: Proceedings of the RESNA '99 Annual Conference, Long Beach, California, USA, 25-29 June, pp.40-42.

Hine, N., Arnott, J.L., 2002. Assistive social interaction for non-speaking people living in the community. In:

Proceedings of the 5th International ACM Conference on Assistive Technologies (ASSETS 2002), Edinburgh, Scotland, UK, 8-10 July, pp.162-169.

Hine, N., Arnott, J.L, Smith, D., 2003. Design issues encountered in the development of a mobile multimedia augmentative communication service. Universal Access in the Information Society 2 (3), 255-264.

Hoag, L.A., Bedrosian, J.L. McCoy, K.F., Johnson, D.E., 2004. Trade-offs between informativeness and speed of message delivery in augmentative and alternative communication. Journal of Speech, Language, and Hearing Research 47 (6), 1270-1285.

Hoag, L.A., Bedrosian, J.L. McCoy, K.F., Johnson, D.E., 2008. Hierarchy of conversational rule violations involving utterance-based augmentative and alternative communication systems. Augmentative and Alternative Communication 24 (2), 149-161.

Hofer, G.O., Richmond, K., Clark, R.A.J., 2005. Informed blending of databases for emotional speech synthesis. In: Proceedings of Interspeech-2005, Lisbon, Portugal, 4-8 September, pp.501-504.

Hutchby, I., Wooffitt, R., 1988. Conversation analysis. Polity Press, Cambridge, UK.

Iwabuchi, M., Alm, N., Andreasen, P., Nakamura, K., Arnott, J.L., 2001. A rapid multi-lingual communicator for non-speaking people and others. In: Proceedings of the IEEE Conference on Systems Man & Cybernetics, Tucson, Arizona, USA, 7-10 October, pp.238-243.

Knapp, M., Hart, R., Friedrich, G., Shulman, G., 1973. The rhetoric of goodbye: verbal and non-verbal correlates of human leave-taking. Speech Monographs, 40, 182-198. Reprinted in: B. Morse & L. Phelps (Eds.), (1980). Interpersonal communication - a relational perspective. Burgess Publishing Company, Minneapolis.

Kraat, A., 1985. Communication interaction between aided and natural speakers: A state of the art report. Canadian Rehabilitation Council for the Disabled, Toronto.

Krivonos, P., Knapp, M., 1975. Initiating communication: What do you say when you say hello? Central States Speech Journal, 26, 2. Reprinted in: B. Morse & L. Phelps (Eds.), (1980), Interpersonal communication - a relational perspective. Burgess Publishing Company, Minneapolis.

Langer, S., Hickey, M., 1998. Using semantic lexicons for full text message retrieval in a communication aid. Journal of Natural Language Engineering 4 (1), 41-56.

Laver, J., 1981. Linguistic routines and politeness in greeting and parting. In: F. Coulmas (Ed.), Conversational routine - explorations in standardized communication situations and pre-patterned speech. Mouton, The Hague.

Lerner, G.H., (ed.) 2004. Conversation analysis: Studies from the first generation. John Benjamins Publishing, Philadelphia.

Levinson, S.C., 1983. Pragmatics, Cambridge University Press, Cambridge.

Light, J., 1988. Interaction involving individuals using augmentative and alternative communication systems: state of the art and future directions. Augmentative and Alternative Communication 4 (2), 66-82.

Lunn, J., Todman, J., File, P., Coles, E., 2004. Making contact in the workplace. Communication Matters 18 (1), 25-28.

Luo, F., Higginbotham, D.J., Lesher, G., 2007. AAC Webcrawler: Enhanced augmentative communication. Presented at the CSUN Annual International Conference on Technology and Persons with Disabilities, Los Angeles, California, USA, 19-24 March.

McCarthy, J., Light, J., 2005. Attitudes toward individuals who use augmentative and alternative communication: research review. Augmentative and Alternative Communication 21 (1), 41-55.

McCoy, K.F., Bedrosian, J.L., Hoag, L.A., 2010a. Implications of pragmatic and cognitive theories on the design of utterance-based AAC systems. In: Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies, Los Angeles, USA, 5th June, pp.19-27.

McCoy, K.F., Bedrosian, J.L., Hoag, L.A., 2010b. Integrating pragmatic theory and empirical evidence into utterance-based system design. In: Proceedings of the 14th Biennial Conference of the International Society for Augmentative & Alternative Communication (ISAAC 2010), Barcelona, Spain, 24-29 July, Paper O.330.

McCoy, K.F., Bedrosian, J.L., Hoag, L.A., Johnson, D.E., 2007. Brevity and speed of message delivery trade-offs in augmentative and alternative communication. Augmentative and Alternative Communication 23 (1), 76-88.

McCoy, K.F., Pennington, C.A., Badman, A.L., 1998. Compansion: From research prototype to practical integration. Natural Language Engineering 4 (1), 73-95.

McKinlay, A., Moore, M., Arnott, J.L., Beattie, W., 1993. Computer-mediated collaborative writing techniques and the disabled. In: Proceedings of the Second European Conference on the Advancement of Rehabilitation Technology (ECART 2), Stockholm, Sweden, 26-28 May, Paper 12.2.

Morris, C., Newell, A., Booth, L., Arnott, J.L., 1991. Syntax PAL – a system to improve the syntax of those with language dysfunction. In: Proceedings of the RESNA 14th Annual Conference (RESNA '91), Kansas City, Missouri, USA, 21-26 June 1991, pp.105-106.

Murray, I.R., Arnott, J.L., 1995. Implementation and testing of a system for producing emotion-by-rule in synthetic speech. Speech Communication 16 (4), 269-390.

Murray, I.R., Arnott, J.L., 2008. Applying an analysis of acted vocal emotions to improve the simulation of synthetic speech. Computer Speech & Language 22 (2), 107-129.

Murray, I.R., Arnott, J.L., Alm, N., Newell, A.F., 1991. A communication system for the disabled with emotional synthetic speech produced by rule. In: Proceedings of the Second European Conference on Speech Communication and Technology (Eurospeech 91), Genove, Italy, 24-26 September, pp.65-68.

Nelson, K., 1991. Remembering and telling: a developmental story. Journal of Narrative and Life History 11 (2-3), 109-127.

Newell, A.F., Langer, S., Hickey, M., 1998. The rôle of natural language processing in alternative and augmentative communication. Natural Language Engineering 4 (1), 1-16.

O'Mara, D., Waller, A., 2003. What do you get when you cross a communication aid with a riddle? The Psychologist 16 (2), 78-80.

Patel, R., Pilato, S., Roy, D., 2004. Beyond linear syntax: An image-oriented communication aid. Assistive Technology Outcomes and Benefits 1 (1), 57-66.

Patel, R., Schooley, K., Radhakrishnan, R., 2006. Comparison of semantic versus syntactic message formulation: A pilot study. Assistive Technology Outcomes and Benefits 3 (1), 96-109.

Patel, R., Radhakrishnan, R., 2007. Enhancing access to situational vocabulary by leveraging geographic context. Assistive Technology Outcomes and Benefits 4 (1), 99-114.

Patel, R., 2010. Leveraging semantic frames and serial icon presentation for message construction. In: Proceedings of the 14th Biennial Conference of the International Society for Augmentative & Alternative Communication (ISAAC 2010), Barcelona, Spain, 24-29 July, Paper O.325.

Polkinghorne, D., 1991. Narrative and self-concept. Journal of Narrative and Life History 1, 135-153.

Pullin, G., Cook, A., 2010. Six speaking chairs (not directly) for people who cannot speak. Interactions 17 (5), September/October, 38-42.

Reiter, E., Turner, R., Alm, N., Black, R., Dempster, M., Waller, A., 2009. Using NLG to help language-impaired users tell stories and participate in social dialogues. In: Proceedings of the 12th European Workshop on Natural Language Generation (Association for Computational Linguistics), Athens, Greece, 30-31 March, pp.1-8.

Roark, B., Fowler, A., Sproat, R., Gibbons, C., Fried-Oken, M., 2011. Towards technology-assisted co-construction with communication partners. In: Proc. of the Second Workshop on Speech & Language Processing for Assistive Technologies (SLPAT 2011), Edinburgh, Scotland, UK, 30th July, pp.22-31.

Rosner, M., 1981. Three strategic goals employed in conversational openings. Working Paper No. 46. Fondazione Dalle Molle, Geneva Institute pour les Etudes Semantiques et Cognitives, Geneva.

Sacks, H., 1995. Lectures on conversation. Blackwell, Oxford.

Sacks, H., Schegloff, E.A., Jefferson, G., 1974. A simplest systematics for the organization of turn-taking for conversation. Language 50 (4), 696-735.

Schegloff, E.A., 2007. Sequence organization in interaction: a primer in conversation analysis. Cambridge University Press, Cambridge.

Schegloff, E.A., Sacks, H., 1973. Opening up closings. Semiotica, VIII, 4, 289-327. Abridged version in R. Turner (Ed.), (1974). Ethnomethodology (pp.233-264). Penguin, Harmondsworth, Middlesex.

Schröder, M., 2009. Expressive speech synthesis: Past, present, and possible futures. In: J. Tao and T. Tan (Eds.), Affective Information Processing. Springer, London.

Searle, J., 1969. Speech acts. Cambridge University Press, Cambridge.

Swiffin, A.L., Arnott, J.L., Newell, A.F., 1987a. The use of syntax in a predictive communication aid for the physically handicapped. In: Proceedings of the 10th Annual Conference on Rehabilitation Technology (RESNA'87), San Jose, California, USA, 19-23 June, pp.124-126.

Swiffin, A.L., Arnott, J.L., Pickering, J.A., Newell, A., 1987b. Adaptive and predictive techniques in a communication prosthesis. Augmentative and Alternative Communication 3 (4), 181-191.

Swiffin, A.L., Pickering, J.A., Arnott, J.L., Newell, A.F., 1985. PAL: An effort efficient portable communication aid and keyboard emulator. In: Proceedings of the 8th Annual Conference on Rehabilitation Technology (RESNA'85), Memphis, Tennessee, USA, 24-28 June, pp.197-199.

Tannen, D., 1984. Conversational style: Analyzing talk among friends. Ablex, New Jersey.

Todman, J., 1999. The use of stored text for socially effective conversation. In: Proceedings of the RESNA'99 Annual Conference, Long Beach, California, USA, 25-29 June, pp.16-18.

Todman, J., Alm, N., 2003. Modelling conversational pragmatics in communication aids. Journal of Pragmatics 35 (4), 523-538.

Todman, J., Alm, N., Elder, L., 1994. Computer-aided conversation: a prototype system for nonspeaking people with physical disabilities. Applied Psycholinguistics 15 (1), 45-73.

Todman, J., Alm, N., File, P., 1999. Modelling pragmatics in AAC. In: F.T. Loncke, J. Clibbens, H.H. Arvidson & L.L. Lloyd (Eds.). Augmentative and alternative communication: new directions in research and practice. Whurr, London.

Todman, J., Alm, N., Higginbotham, J., File, P., 2008. Whole utterance approaches in AAC. Augmentative and Alternative Communication 24 (3), 235-254.

Todman, J., Elder, L., Alm, N., 1995. Evaluation of the content of computer-aided conversations. Augmentative and Alternative Communication 11 (4), 229-234.

Vanderheiden, G.C., 1983. Non-conversational communication technology needs of individuals with handicaps. Rehabilitation World 7 (2), 8-12.

Vanderheyden, P.B., 1995. An augmentative communication interface based on conversational schemata. In: Proceedings of the IJCAI '95 Workshop on Developing AI Applications for Disabled People, Montreal, Quebec, Canada, pp.203-212.

Vanderheyden, P.B., Demasco, P.W., McCoy, K.F., Pennington, C.A., 1996. A preliminary study into schema-based access and organization of reusable text in AAC. In: Proceedings of the RESNA '96 Annual Conference, Salt Lake City, Utah, USA, 7-12 June, pp.59-61.

Waller, A., 2006. Communication access to conversational narrative. Topics in Language Disorders 26 (3), 221-239.

Waller, A., Alm, N., Newell, A.F., 1990. Aided communication using semantically linked text modules. In: Proceedings of the RESNA 13th Annual Conference, Washington, DC, USA, 15-20 June, pp.177-178.

Waller, A., Dennis, F., Brodie, J., Cairns, A.Y., 1998. Evaluating the use of TalksBac, a predictive communication device for non-fluent adults with aphasia. International Journal of Language and Communication Disorders 33 (1), 45-70.

Waller, A., O'Mara, D., Tait, L. Booth, L., Brophy-Arnott, B., Hood, H., 2001. Using written stories to support the use of narrative in conversational interactions: case study. Augmentative and Alternative Communication 17 (4), 221-232.

Waller, A., Black, R., O'Mara, D.A., Pain, H., Ritchie, G., Manurung, R., 2009. Evaluating the STANDUP pun generating software with children with cerebral palsy. ACM Transactions on Accessible Computing 1 (3), Article 16.

Wisenburn, B., Higginbotham, D.J., 2008. An AAC application using speaking partner speech recognition to automatically produce contextually relevant utterances: Objective results. Augmentative and Alternative Communication 24 (2), 100-109.

Wisenburn, B., Higginbotham, D.J., 2009. Participant evaluations of rate and communication efficacy of an AAC application using natural language processing. Augmentative and Alternative Communication 25 (2), 78-89.

Wittgenstein, L., 1953. Philosophical Investigations. Basil Blackwell, Oxford.

Wray, A., 2001. Formulaic language and the lexicon. Cambridge University Press, Cambridge.

Yang, G., McCoy, K., Demasco, P., 1990. Word prediction using a systemic tree adjoining grammar. In: Proceedings of the RESNA 13th Annual Conference, Washington, DC, USA, 15-20 June, pp.185-186.

Yngve, V., 1970. On getting a word in edgewise. In: Papers from the Sixth Regional Meeting of the Chicago Linguistic Society, Vol.6, pp.567-578. Chicago Linguistic Society, Chicago.