Automatic generation of conversational utterances and narrative for Augmentative and Alternative Communication: a prototype system

Martin Dempster & Norman Alm
School of Computing
University of Dundee
Dundee, Scotland, DD1 4HN, UK
m.k.dempster@dundee.ac.uk
nalm@computing.dundee.ac.uk

Ehud Reiter
Computer Science Department
University of Aberdeen
Aberdeen, Scotland, AB24 3UE, UK
e.reiter@abdn.ac.uk

Abstract

We detail the design, development and evaluation of Augmentative and Alternative Communication (AAC) software which encourages rapid conversational interaction. The system uses Natural Language Generation (NLG) technology to automatically generate conversational utterances from a domain knowledge base modelled from content suggested by a small AAC user group. Findings from this work are presented along with a discussion about how NLG might be successfully applied to conversational AAC systems in the future.

1 Introduction

Augmentative and Alternative Communication (AAC) systems assist non-speaking people communicate. Reasons for lack of speech are varied and can be complex, but they are typically related to some profound cognitive and/or motor impairment.

Most AAC systems are computer based, utilize synthesized speech output and employ a *phrase-construction* approach to input. This approach requires the user to construct the majority of their utterances live during conversation. Undoubtedly this facilitated communication is hugely important to those without natural speech. However, this process is often unacceptably slow and can lead to problematic and stilted interactions, mostly due to the rapid nature of unimpeded face-to-face communication.

Previous work has shown that it is possible to hold mutually rewarding conversations using wholly pre-stored material, known as the *phrase-storage* approach. Utterances are authored ahead of time and can be selected and output immediately leading to quicker communication rates. However, this approach suffers from several drawbacks which may have affected its more general adoption.

Furthermore, Natural Language Processing (NLP) technology has proven to be a fruitful line of inquiry within the field. It has offered a powerful means to improve system productivity and usability. We are currently investigating how Natural Language Generation (NLG) might be applied in a useful way within an AAC device geared towards fast-paced and rewarding social interactions. It is hoped that the linguistic control and automaticity offered by NLG may go some way towards addressing the previous criticisms of pre-stored material regarding its inflexibility and cost in effort.

2 Background

2.1 Limitations of current AAC

High-tech AAC systems typically augment communication for non-speaking people by allowing live message construction through some orthographic means. Completed messages are generally sent to a speech synthesis engine for output during communication with others. Many people who require AAC have associated physical disabilities which reduce the speed achievable using input methods such as keyboards, pointing devices or touch-screens. The rate achievable using most commercial AAC systems is highly dependent on the nature of the user's disabilities but a generally accepted figure is in the region of 2-15 words per minute (Higginbotham, Shane et al. 2007), at least an order of magnitude slower than most natural speakers.

This relatively slow rate of input is a crucial factor in some of the issues that arise in AAC-facilitated communication. Because of the effort and time required to create utterances, the user may not be able to construct messages quickly enough to take active roles in fast

Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies, pages 10–18,
Los Angeles, California, June 2010.

paced conversations. As a result users may become passive while also typically using a smaller communicative repertoire than natural speakers (Light 1988).

Narrative, an important type of interpersonal communication, is not well handled in most communication aids (Waller 1992). Delayed response and slow rate of aided-communication are correlated with higher incidence of breakdown in communication and lesser perceptions of the AAC user (Todman and Rzepecka 2003; McCarthy and Light 2005). This is primarily due to conflict between the relatively long time necessary to formulate an utterance and the fast paced nature of conversation.

These problems are particularly critical in social contexts. AAC users typically have small social circles and are dependent on contact with families and carers. They often lack self-esteem and have negative self-image. As a result, developing new relationships and experiencing new things can be difficult, despite being a major priority in their lives (Datillo, Estrella et al. 2007).

Some work has suggested that the use of pre-stored conversational material based on conversation models could help increase communication rate and conversation quality. Alm (1988) showed that it is possible to successfully model short ‘chat’ conversations involving greetings, personal enquiries and small-talk. Furthermore, the TALK system allowed a user to pre-store a large volume of material on specific topics so that whole utterances could be selected and output. The system also made heavy use of *quick-fire* phrases, classes of regularly used utterance which could be accessed quickly, and showed that communication using solely pre-stored material was viable (Todman and Alm 2003).

Despite encouraging results and the development of a commercial product, the *phrase-storage* approach to social communication has not gained wide popularity. The reasons for this are complex, but include: the relative inflexibility of pre-stored material; the costs associated with authoring the material and keeping the material up-to-date; and the vastly different nature of the approach and different training requirements necessary to achieve success.

2.2 The role of NLP in AAC

NLP technology has provided many benefits to AAC system designers. Possibly the first technology to be included in many commercial systems to date was word prediction and completion. There have also been many research prototypes exploring the applicability of more emerging technologies such as named entity recognition from synthesized speech (Wisenburn and Higginbotham 2008), the generation of well-formed utterances from telegraphic input (McCoy, Pennington et al. 1998) and the automatic identification of contextual vocabulary from the web (Higginbotham, Bisantz et al. 2008). Netzer and Elhadad (2006) used NLG to allow the semantic authoring of utterances.

However, NLG, in the sense of data-to-text (Reiter and Dale 2000), has had limited application within AAC thus far, although Reiter et al. (2009) showed it is possible to generate stories from sensor data which allow a child using AAC to tell others about their day at school.

2.3 System Rationale

This project is exploring the use of NLG to produce conversational utterances in AAC systems designed for social interaction. At the outset it was hoped that using NLG might address some of the difficulties observed in pre-storage systems. For instance, the generation component could theoretically produce a range of utterances and speech act types automatically from the same underlying data and adapt these somewhat to the interactional context. Using NLG would also have the benefit of offering control over the well-formedness of the output, an important consideration given the difficulty some AAC users have in achieving literacy (Sandberg and Hjelmquist 1997). The fact that the system has an inherent awareness of the semantic content of the linguistic output, rather than simply being stored as canned text, is also a potential benefit. In other words, NLG might offer a level of automaticity and flexibility that traditional pre-storage systems cannot offer, as well as potentially reducing the level of pre-authoring required from the user.

3 System Development

3.1 User-centered methodology

To try to assess how useful NLG could be in this context we adopted a user-centered approach to the design of the system. A group of 3 AAC users has been recruited, all of whom currently use some form of high-tech AAC. Literacy amongst the group is varied. Two of the individuals use the alphabetic keyboard-based Lightwriter communication device currently, and have normal cognitive and visual-perceptive skills. All of the users have cerebral palsy and dysarthria, and have been involved in previous software evaluations.

Weekly or twice weekly sessions were held with each user for several months while the software was being produced. Sessions consisted of various activities:

discussion about the user's ideas for the software and technology; the identification of topics and collation of input data to the system; demonstrations of the new features or changes since the last session; system training; and *dry-run* test conversations between the investigators and the users.

3.2 System Architecture

A growing line of inquiry in the NLG community is the generation of language from ontologies (Mellish and Sun 2005). An ontology is a logical and hierarchical model of the different concepts and the nature of relationships between concepts in a particular domain. These concepts and relationships can be mapped onto linguistic constructs to allow for the production of natural language descriptions (Karakatsiotis, Galanis et al. 2008) of parts of the ontology.

In the case of our system, we are trying to model conversational topics that would be of interest in social conversation between users of the system and their co-conversationalists. The current categories of topic we are experimenting with include travelling, listening to music, watching films and attending concerts. Many categories are based on a simple event model which defines the basic characteristics common to all events, such as a time of occurrence (see Fig.1). We have also included concepts such as Person and Place which are associated with events to form a logical model of a particular event type.

A separate file is created unique to each user which is linked to the original model. This is filled with *individuals* consisting of data from the user. In other words, rather than defining the concept of an event as we did with the original ontology, here we are creating a description of an actual event and any other details, such as people or places, associated with it. We have defined our ontology in OWL, a standard language for the definition of ontologies, and each piece of knowledge is effectively stored as a RDF Triple consisting of a subject, predicate and object.

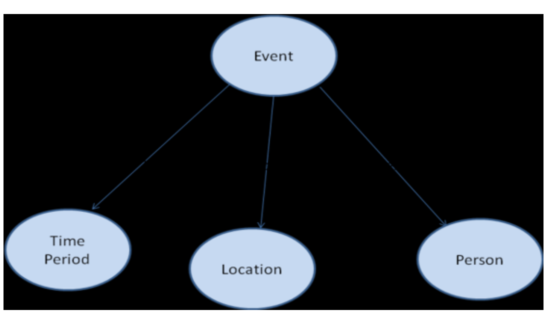

Figure 1: The abstract event model

The user's knowledge base is turned into useful conversational utterances through a template-driven utterance generation system (e.g. Van Deemter, Krahmer et al. 2005). A large set of templates has been authored, using the SimpleNLG programming interface, which turn data from the onotlogy into natural language utterances. The templates are created as *concrete syntax trees* containing unspecified 'slots' and parameters (See Fig.2). These syntax trees map out the syntactic structure of the template), and are linked to a particular class in the ontology so that only appropriate templates are applied to each *individual*. Slots are used to add contextually relevant clauses to our utterances. For example, a template might contain a *'time'* slot, the contents of which are derived from the time of the event in question. For instance, the slot might be filled with "next Tuesday evening", "a month ago" or "this morning" depending on the context. Example parameters include the *tense* with which the utterance should be generated, and whether a pronoun or full noun phrase should be used to refer to the subject of the utterance.

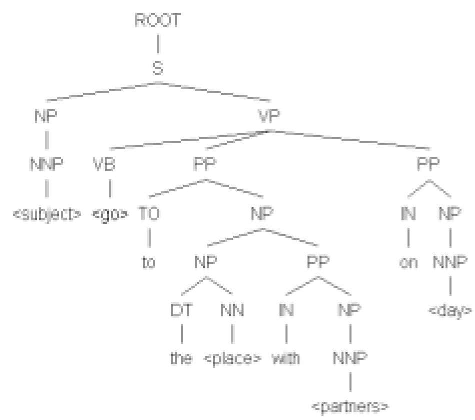

Figure 2: An example syntax tree with empty slots

In addition to the language produced from the model and knowledge base we have included the ability to add canned text phrases to each individual.

This is necessary because there may be things that you wish to be able say about a topic which it is not feasible to model. Because we have a fairly diverse set of topics it is simply not possible to model all aspects of these topics in a reasonable time. There is effectively a trade-off between complexity of the model and how maintainable and representative it is. A more complex model will lead to more expressive generated language, but will cost a great deal more to design and maintain. In the case of our system, a 'lowest common denominator' domain model combined with additional canned text has proven to be a relatively straightforward and inexpensive design.

The system has also been designed to learn over time the sequences of utterances a user selects and suggest next moves based on past behavior. The system does this by maintaining a *directional weighted graph* which records sequences of utterances as they are used. The graph works by recording each individual utterance as a node in the graph and creating relationships between these nodes as they occur. The more often two utterances appear in sequence the higher the value given to the edge between the two corresponding nodes.

3.3 Conversation model and interface issues

Perhaps the most challenging aspect of taking the system from initial concept to working prototype has been finding the most effective way of interfacing the technology. We have found that due to the complexity of the underlying technology, reaching the stage where generated utterances are both useful and accessible to the user during conversation has required careful consideration and the trialing of several approaches with the user group.

It was envisaged that the generative power of NLG would be its most powerful benefit. The system could realize the same piece of data as numerous speech act types and, within a speech type, in several different phrasings. This offered the ability to counteract the inflexibility and uniformity of pre-stored utterances somewhat. However, we have had mixed success in achieving this goal as it has proven difficult to find an effective way to interface this enhanced choice and variety to the user. If there is a large volume of generated utterances available to choose from we must provide an efficient means by which the material is presented or organized so that the desired utterance can be located quickly. If a large choice results in a delayed selection and thus conversational turn, we may then lose any rate and speed of response benefits which would negate the need to use pre-stored and generated material at all.

To address this problem, we attempted to design a conversational model which controlled the generation of utterances so that only the utterances deemed most likely were presented to the user, thus reducing the cognitive load required to search through a large set. This was done using a basic system where the templates were tagged according to where it might be most likely to be used in a conversation on a topic. For instance, a template might produce a pre-sequence, an introduction, elaboration or concluding remark, or it may produce a interrogative. With the addition of historical sequential moves from our directional graph we could begin to present subsets of utterances to the user according to where they were in topic development.

Another approach trialed was inspired by the Gricean maxim of quantity. Each template contains meta-data about the information it expresses. For each generated utterance selected, we can 'rule out' further generation of the same information. This is based on the assumption that speakers will generally avoid repetition. We have found that this technique provides a useful way of supporting discourse coherence within conversations.

Finally, using the logical model of topics we have created, it is possible to support and model stepwise topic progression. We can suggest, based on the model and the user knowledge base, other topics linked to the one currently selected. For example, if we were talking about an upcoming holiday to London with a friend called Bob we may want to the change topical perspective to related aspects of the trip. We might want to talk about London as a place, Bob as a friend, and other trips we have taken with Bob or to London. Because these concepts are all distinct within the model, they each have their own set of associated templates and result in sets of candidate utterances with differing perspectives. Navigating to related topics in this manner should be quicker since related topics do not have to be located manually. Although the users are still being trained in this approach to topic change, early evaluations are promising. It enables a 'one-click' transition to related topics, allowing the user to elaborate on certain aspects of a previous topic and respond quickly to questions from their conversational partners.

Building on the last two mechanisms, we can also generate bridging phrases which allow for more cohesive changes in topic. This allows for a more eloquent transition to a new topic and also aids the discourse coherence.

All of these approaches in fact belie, to some degree, the complexity of conversation. By its very nature, conversation is unpredictable, and the purpose and meaning of sequential moves are highly dependent on their context (Clark 1996). However, any form of context identification, such as speech recognition (Wisenburn and Higginbotham 2008), is likely to present a major technical challenge in any production AAC system at the current time. The above are simply at attempt to model, using the NLP/AI techniques available, aspects of communication process, to show the potential benefits when using NLG-produced utterances rather than simple canned text utterances.

Application of some of the above techniques resulted in a highly fluid interface in which the utterances displayed changed rapidly according to the conversation model. This presented a major challenge to users learning the system, with all displaying a strong preference for a static interface where the same utterances could be found in the same location each time they were desired.

	UTTERANCE	USER SELECTION
A:	**Hi Robert**	**[GREET]**
B:	Oh, Hi. Nice to see you.	
A:	**And you.**	**[GREET]**
A:	**How's it going?**	**[INTRO]**
B:	Fine. And you?	
A:	**Not too bad.**	**[INTRO]**
B:	So you been keeping busy?	
A:	**Yeah**	**[YES]**
A:	**I certainly have!**	**[YES]**
A:	**I was out at a concert on Thursday night. (G)**	**[GIGS] [Select 'Martin Taylor']**
B:	Great. Who did you go to see?	
A:	**Have you heard of Martin Taylor? (G)**	**[ARTIST] [Select 'Martin Taylor']**
B:	No.....I don't think so.	
A:	**He is a Jazz guitarist. (G)**	**[Select 'Martin Taylor']**
B:	Oh, great. I like jazz music.	
A:	**Me too.**	**[AGREE]**
B:	So how was the concert?	
A:	**It was really good. (G)**	**[GIGS] [Select 'Martin Taylor']**
A:	**John and David came with me. (G)**	
A:	**We all enjoyed it. (C)**	
A:	**We had a bit of an interesting journey home because it was snowing heavily, but we made it back safe. (C)**	
B:	Well that's good news. Where was the concert?	
A:	**It was at the Tron Theatre in Glasgow. (G) I've been to Glasgow a few times lately. [G]**	**[GIGS] [Select 'Martin Taylor'] [Select 'Glasgow']**
A:	**Anyway, I best be getting on.**	**[WRAP UP]**
A:	**It was great talking to you.**	**[WRAP UP]**
B:	Yes, likewise.	
B:	See you soon.	
A:	**OK. Cheerio.**	**[FINISH]**
B:	Bye	

Table 1: An example conversation produced using the system. Speaker A is the user and speaker B is an unaided speaker. The right-hand column shows the interface selections necessary prior to selecting the utterance from a set of possibilities. The marker G represents a generated utterance, C represents canned text. The remainder are *quick-fire* utterances.

We believe this does not suggest that use of such conversational models and semantic processing is not feasible, but simply that in the scope of the current work it has not been possible to fully evaluate their potential. Thus we have chosen to generate candidate phrases in a static manner without the predictive aspects described above. These changes have allowed for quicker achievement of proficiency and have lowered the cognitive effort required to navigate the interface.

In the latest version of the software, we have defined a set of templates for each topic which when realized in series produce a coherent narrative. They can still be selected individually by the user for output, so they retain ultimate control of what is said, but the utterances are presented in a natural order. This means that the user can easily make use of the utterances as a narrative or can choose according to the particular situation and context. Any interrogative templates are displayed in a different part of the interface. We have set up a two column display so that interrogatives and other statements are clearly delineated.

This approach has had very promising results as we have found that users no longer have to search through a list of suggestions which changes after each conversational turn. They can also use the structured nature of the generated utterances to confidently introduce the different topics in conversation. We are finding some evidence of increased self-selection at the end of their current turn as the user is easily able to continue their narrative automatically without having to worry about the location of their next turn in the interface. There is some other evidence of this structured application of NLG to narrative as being a promising area (Reiter, Turner et al. 2009).

We also believe that the passivity and lack of initiation observed in AAC users could be positively addressed if AAC systems can better support a more varied communicative repertoire and suitable training is administered to show users how to confidently use these different constructs (e.g. Todman 2000). Early training sessions with our user group have again proved positive with increased use of the trained features and interaction styles.

Figure 3 - System interface

3.4 Authoring user content

Currently we have not managed to produce a tool that the user can use to update their knowledge base themselves. The ontology editing tool used in the program, Protégé, is a free academic software package designed for knowledge engineers and thus has a high degree of internal complexity and takes time to learn. It is also not a particularly accessible piece of software.

We have worked with the users to build up their knowledge bases over a series of meetings by allowing them to suggest individuals to add while entering the details for them into the system. The process of defining new individual is very quick, usually requiring the input of just a few words and selection of the associated individuals. However, one of the main criticisms on the part of the users is that for the system to be useful in the long term, it must be kept up to date, as old material will quickly become less relevant and useful in less frequent situations. For this reason it is critical to the success of any NLG-driven communication system that the data input is as simple and seamless as possible.

We have shown in our system that it is possible to get some limited data automatically from online sources, rather than having to input it manually. Many web services are being made available which enable programmers to access data from online services in their applications. For instance, both Amazon and YouTube have their own APIs which allow 3rd party applications to request content information from these services.

The notion of the semantic web is also related to this. There is a large effort underway to define how we might structure and link information on the web in such a way that more of it can be processed automatically by computers and made available in interchangeable formats.

Shared data and semantic web technologies such as these operate on the same premise as our proposed communication system in that they map out the basic vocabulary required to describe a domain, and allow people describe aspects of the domain in these terms.

We have used an API provided by social music website Last.fm to show that it is possible to create relevant conversational utterances without any authoring requirement whatsoever. By supplying the users Last.fm username, we can use a web service supplied by the site to query the user's recent activity, for instance the songs they have listened to, songs rated highly or events which they have signed up to attend. Because the output from these services comes as structured XML document we can simply map it's schema onto our own vocabulary and feed the appropriate data to our templates to produce utterances.

If web services are to be used we must have an equivalent local vocabulary to which we can map the data returned from any queries we send the service. However, in the case of semantic web sources, for example the FOAF (friend-of-a-friend) vocabulary (Brickley) describing online social networks, the process is simpler as we can simply use the pre-existing vocabulary standard ourselves rather than having to develop our own. Despite the semantic web being in its infancy, the notion of shared data is growing in popularity and many popular websites and organizations are providing access to their information in a structured way.

One problem with using these types of data acquisition methods for our purposes is that the data is largely generic and any personal opinion or evaluative information personal to the user is limited. In some cases we may be able to query the data source for a rating awarded to a particular piece of content, for instance the star rating system on YouTube, but it is not clear how expressive the produced language will be since the process is likely to be a simple mapping from the rating to a suitable adjective. As in our system, the potential usefulness of the generated language is likely to be increased if it is possible for the user to annotate the topics with their own canned text expressions and evaluations. This will enable the system to express more of the individual's personality and opinions.

We believe this is an area of great interest for AAC. There is growing evidence of the importance of the internet in the lives of disabled people, particularly its role as a communication medium for people with communication impairments (Cohen 1999). By harnessing the large volumes of data created when using modern hardware and software systems and transforming it into useful utterances, we can begin to address one of the main criticisms of whole utterance approaches to AAC since there would be no authoring requirement on the part of the users. This is certainly by no means a simple process and this approach will require further investigation, but as semantic web technologies reach maturity and gain wider adoption it should be clearer what the potential of the technology is.

4 Formal Evaluation Methodology

In our evaluations so far, we have concentrated on training the users in its operation, updating conversational material and implementing changes based on the user feedback. We have recently begun testing the system in real conversational encounters and the results have been promising. We have found it is possible to hold pleasing conversations lasting up to 20 minutes with unfamiliar partners, with the aided communicator achieving a rate of upwards of 40 wpm.

There also seems to be higher incidences of initiations on the part of the user, with them making good use of both the scripted NLG material, the quick fire phrases and their own pre-stored material. The topic progression feature is currently being underused but subjects are responding well to training sessions on how to incorporate this to reduce their response time and expand on topics to extend the amount they are able to say.

Formal evaluations are now being undertaken. An AB multiple-baseline study design is being conducted in which each aided communicator has a series of conversations with 12 unknown and unaided conversation partners. In the A condition, the aided participants use their existing AAC system, while in the B condition they use our prototype system. Each conversation will be limited to approximately 10 minutes, and the sessions will be split across a three non-consecutive days to avoid user fatigue.

There will be at least 3 conversations in both the A and B conditions, and the intervention point will be randomized across the remaining 6 conversations to allow for valid inferences to be made despite the small *n* value (Todman and Dugard 2001). This also reduces the bias introduced by any training effects and avoids the need to use a response-guided intervention after baseline performance has been established. The difficulty and expense of recruiting large numbers of subjects in AAC studies is a known problem (Higginbotham 1995) and therefore any findings from quantitative analysis performed cannot be generalized across the AAC population. However, we expect to be able to achieve a p value using the randomization design of <0.05 so the results should at least be internally robust and give a good indication of whether further investigation is warranted.

The conversations will be audio-recorded and analyzed for a number of metrics. Primarily we are interested in measuring the rate at which people are able to communicate using the new system as this seems to be one of the clearest indicators of success when evaluating a new AAC intervention. We expect to the effect size observed across the conditions to be large.

We are also particularly interested whether it is possible to use automatically generated material while maintaining or enhancing the enjoyment and quality of the encounter for all participants. It is still unclear how acceptable generated material will be to the user so we will measure the relative frequency of generated and canned-text utterances.

In previous studies it has also been shown that the use of a whole-utterance approach can change the dynamics of communication, such as relative speech act distribution and number and type of initiation, so we are interested to see how the availability generated material might impact this and what role it might play. A coding schema based on Wang (2007) will be used to categorize the utterances used.

We are also asking the aided and unaided conversation partners to complete questionnaires regarding various subjective ratings of the interactions and, in the case of the unaided speakers, impressions of the aided communicator. The questions will be based on a re-evaluation of those suggested by Todman (2000) and answers will be requested on a 7-point rating scale. Previous work has shown that quicker, flowing interactions with less breakdowns or delays can lead to more rewarding interactions for both participants. We expect to observe these effects in our system but it's as yet unclear what impact the automatically generated phrases will have, if any, on perceptions of the user.

Although the relatively small number of participants means it is unlikely that we will be able to make robust inferences from this data, we hope that results will be indicative of the naturalness and acceptability of automatically generated utterances.

5 Discussion & Future Plans

One of the primary reasons that AAC systems featuring NLP technology prove useful is that they go some way to leveling the playing field for many users. They have the potential to support the user in ways which reduce the effort required to communicate yet may improve the quality of the communication. There are many NLP technologies, such as NLG, that deserve further attention within the field of AAC to determine what they can offer.

Although our system has shown some encouraging preliminary results there are still many unanswered questions with regards to the role NLG can play. For example, it is not clear how appealing NLG utterances are to use. Given that the user has not authored the form of the utterances themselves there is an argument that using them may feel unnatural. After the formal evaluations we should be able provide analysis indicating whether NLG phrases are being used and in which situations they are proving most useful.

One of the most challenging aspects of designing the system was the HCI challenge of incorporating some of these technologies. While it is obvious to the user that phrases are being generated automatically, and that these phrases are generated when a topic is selected, it is still important to note that the technologies have been intentionally kept largely transparent to the user. When using a communication system, the most important thing is the ability to say what you want to say, but is not yet clear whether the technical nature of the software may be an alienating factor since the user currently has no access to the template construction or domain modeling aspects of the system.

At the current time, the domain modeling and template construction processes are quite complex and expensive. Tools are becoming available, from the NLG community, which go some way to addressing the difficulty of interfacing these types of technology to non-experts (Bilidas, Theologou et al. 2007; Power, Stevens et al. 2009) but these are largely unsolved problems.

Domain modeling itself is problematic in that one persons notion of what defines a particular concept is often different to someone else's. For instance, one person's idea of sport might encompass the sporting activities they take part in, while another person's idea of sport is that which they follow or watch on the television. This has clear implications for the general usability of the system. Using semantic web vocabularies may address this somewhat since they are likely to be more specific to a particular purpose and be more mature and interoperable than the *'home-brew'* domain models we have used for the prototype.

Using whole-utterance approaches to communication clearly requires the adoption of a different mindset. Rather than being able to construct a novel message the user has to '*make do'* with whatever is available in the system. Despite the advantages observed while using such systems, they have still not become generally popular. It is likely that any NLG whole utterance system would similarly not gain immediate acceptance because it is vastly different to other systems and approaches to communication available. To some degree we are asking the user of our NLG system to think in an object orientated manner since they must understand the underlying model and the way the concepts are structured to make the most of the system. Again it is not yet clear how natural this process is and how much training is required to become an expert user of such a system.

However perhaps the major strength of these types of system is the way in which they help scaffold interaction so the AAC user can be much more active in conversation and use an increased repertoire. The design of the software is such that it encourages the use of types of phrases often underused by AAC users, for example, initiations, elaborating moves, questions and the different classes of quick-fire remarks. One interesting question is whether the use of NLG might make it easier to encourage the user to use new types of conversational move. Since no full text-authoring is required the user does not even have feel confident authoring the utterance, it is simply provided and can be used or experimented with. Scaffolding interactions in this way may be one of the most interesting avenues for NLP and AI technologies with AAC in the future.

The architecture of the prototype, although effective, lacks efficiency and may be difficult to reuse. A great deal of work is being done by NLG researchers investigating how NLG architectures might be made more modular and reusable. This is an ongoing problem but it seems sensible to consider how a pipeline architecture (Reiter and Dale 2000) might work in practice for this type of system.

At the moment, the system requires a reasonable level of literacy because the interface is mainly text based. However, semiotic systems are preferred because of the literacy problems observed in many AAC users. It is not clear how NLG may impact on semiotic message construction but systems such as Compansion (McCoy, Pennington et al. 1998) show there may useful applications in this area too.

6 Conclusion

Despite having only been able to perform informal evaluations so far, we believe we have seen some encouraging signals that NLG may have potential as an augmentative communication technology to assist in generating conversational utterances. We believe that the rapid access to well-formed, contextually generated material offered in our system could lead to significant benefits for the AAC user and their interlocutors.

There are further exciting possibilities with regards to the technology, particularly the ability to harvest personal data from the internet and other computer usage so that it can be transformed into useful phrases for inclusion in communication aids. We hope to have a richer set of data and results in the coming months after the system training and formal evaluations have been completed.

7 References

Alm, N. A. (1988). Towards a Conversation Aid for Severely Phisically Disabled Non-Speaking People. Applied Computing Department. Dundee, University Of Dundee. **Doctor Philosophy:** 197.

Bilidas, D., M. Theologou, et al. (2007). Enriching OWL Ontologies with Linguistic and User-related Annotations: the ELEON system. 19th IEEE International Conference on Tools with Artificial Intelligence, IEEE.

Brickley, D. (25.2.10). "FOAF Vocabulary Specification ", from http://xmlns.com/foaf/0.1/

Clark, H. H. (1996). Using Language, Cambridge University Press.

Cohen, K. J. (1999). Using the Internet to Empower Augmented Communicators. CSUN'99.

Datillo, J., G. Estrella, et al. (2007). ""I have chosen to live life abundantly": Perceptions of leisure by adults who use Augmentative and Alternative Communication." Augmentative & Alternative Communication **24**(1): 16-28.

Higginbotham, D. J. (1995). "Use of nondisabled subjects in AAC Research : Confessions of a research infidel." Augmentative and Alternative Communication **11**(1): 2-5.

Higginbotham, D. J., A. M. Bisantz, et al. (2008). "The Effect of Context Priming and Task Type on Augmentative Communication Performance." Augmentative & Alternative Communication.

Higginbotham, D. J., H. Shane, et al. (2007). "Access to AAC: Present, past, and future." Augmentative & Alternative Communication **23**(3): 243-257.

Karakatsiotis, G., D. Galanis, et al. (2008). NaturalOWL: Generating Texts from OWL Ontologies in Protege and in Second Life. 18th European Conference on Artificial Intelligence.

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.

McCarthy, J. and J. Light (2005). "Attitudes towards individuals who use Augmentative and Alternative Communication: Research Review." Augmentative and Alternative Communication **21**(1): 41-55.

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

Mellish, C. and X. Sun (2005). The Semantic Web as a Linguistic Resource: Opportunities for Natural Language Generation. International Conference on theory, practical and application of Artificial Intelligence. M. Bramer, F. Coenen and T. Allen. Cambridge, UK, Springer**:** 77.

Netzer, Y. and M. Elhadad (2006). Using Semantic Authoring for Blissymbols Communication Boards. HLT-2006.

Power, R., R. Stevens, et al. (2009). Editing OWL through generated CNL. Workshop on Controlled Natural Language (CNL'09). Marettimo Island, Italy.

Reiter, E. and R. Dale (2000). Building Natural Language Generation Systems. Cambridge Cambridge University Press.

Reiter, E., R. Turner, et al. (2009). Using NLG to help language-impaired users tell stories and participate in social dialogues. Proceedings of the 12th European Workshop on Natural Language Generation. Athens, Greece, ACL.

Sandberg, A. D. and E. Hjelmquist (1997). " Language and literacy in nonvocal children with cerebral palsy." Reading and Writing **9**(2): 107-133.

Todman, J. (2000). "Rate and quality of conversations using a text-storage AAC system: Single-case training study." Augmentative and Alternative Communication **16**: 164-179.

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

Todman, J. and P. Dugard (2001). Single-case and small-n experimental designs: A practical guide to randomisation tests. Mahwah, NJ, Lawrence Erlbaum Associates.

Todman, J. and H. Rzepecka (2003). "Effect of pre-utterance pause length on perceptions of communicative competence in AAC-aided social conversations." Augmentative and Alternative Communication **19**(4): 222-234.

Van Deemter, K., E. Krahmer, et al. (2005). "Plan-based vs. Template-based NLG: A false opposition?" Computational Linguistics **31**(1).

Waller, A. (1992). Providing Narratives in an Augmentative Communication System. Applied Computing. Dundee, University Of Dundee. **Doctor of Philosophy:** 163.

Wang, Y. (2007). A model of conversational structure for augmentative and alternative communication (AAC) systems. University of Dundee, Unpublished PhD Thesis. **PhD**.

Wisenburn, B. and D. J. Higginbotham (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.