This is a repository copy of *What is the potential for context aware communication aids?*.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/90718/

Version: Accepted Version

**Article:**
Judge, S., Hawley, M.S., Cunningham, S. et al. (1 more author) (2015) What is the potential for context aware communication aids? Journal of Medical Engineering & Technology, 39 (7). pp. 448-453. ISSN 0309-1902

https://doi.org/10.3109/03091902.2015.1088091

---

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher’s website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
What is the Potential for Context Aware Communication Aids?

Simon Judge $^{1,2,3}$, Mark S Hawley$^{1,2,3}$, Stuart Cunningham$^{1,4}$, Andrea Kirton$^{2}$

1. University of Sheffield, Centre for Assistive Technology and Connected Healthcare and School of Health and Related Research
2. Barnsley Hospital, Assistive Technology Team
3. NIHR Devices for Dignity Healthcare Technology Co-operative
4. University of Sheffield, Department of Human Communication Sciences

Author Note

Correspondence Address: School of Health and Related Research (ScHARR), University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

0114 222 0682

Simon.judge@nhs.net; mark.hawley@sheffield.ac.uk; s.cunningham@sheffield.ac.uk; andreakirton@nhs.net

Declaration of interests: None.

This is an Accepted Manuscript of an article published by Taylor & Francis in JMET on 29-09-15, available online: http://www.tandfonline.com/doi/abs/10.3109/03091902.2015.1088091
Abstract

Use of voice output communication aids (VOCAs) can be a very effective strategy to assist people with speech impairments in communicating. Despite this, people who use communication aids often express frustration with VOCAs – desiring devices that are simpler, quicker and more effective to use.

Whilst it is not possible to resolve all these issues with technology, we argue that significant progress can be made. The use of contextual information is one development that could improve the simplicity and effectiveness of communication aid design.

Improving the effectiveness of communication aids, including through the use of context support, is a goal of the NIHR Devices for Dignity Assistive Technology Theme. This discussion paper examines the potential for creating ‘context aware’ communication aids. Three projects in which the authors have been involved are described to illustrate different approaches to the use of contextual information.

Keywords: Communication Aid, AAC, Context, NLG, NLP
What is the Potential for Context Aware Communication Aids?

Augmentative and Alternative Communication (AAC) refers to a range of strategies, techniques and devices that can support the communication of individuals with speech, language and communication impairments. Voice output communication aids (VOCAs) are one AAC strategy often used by individuals to support their communication. VOCAs are devices which take an input from the user and output synthesised (or recorded) speech. Here we are considering the use of AAC for expressive communication (as defined by Tetzchner & Martinsen), although AAC and VOCAs can also be used to support language development.

It can be argued that VOCAs to date have been designed predominately for individuals with no functional speech, but they are also used by those whose speech intelligibility is variable or understandable only to people who know them well (e.g. when used by individuals with severe or moderate dysarthria). VOCAs are used rarely as the only mode of communication of an individual, but tend to be used as part of an overall set of communication strategies.

VOCAs may use text or graphic symbols to represent language on the device and as an input method for the user. Depending on the type of AAC being used, items that can be selected by the user may be letters (or other graphemes), whole words or phrases, or may be other graphic symbols representing words, concepts or utterances.

Individuals using a VOCA also face constraints related to the selection of these items. Those using text representation will need to select around 27-40 items (unless using an ambiguous keyboard), whereas those using graphic symbols may need to select from as many as 4000 items. To navigate this number of items the user is required to make a series of selections.
either as a coded sequence (e.g. navigation through pages) or with an explicit code (e.g. Morse or quartering\textsuperscript{6}).

Individuals use a range of strategies to select items on a VOCA - some may be able to press directly onto a screen or keyboard (known as direct access), whereas others may have difficulties with co-ordination or control and require alternative access devices. These devices can include alternative keyboards (for example those with key ‘guards’), alternative mice (for example, eye gaze or head movement selection), or other controls such as switches. Switches are binary (non-latched) ‘buttons’ that come in a range of types - with varying size, shape, activation force or method. When using a switch, the user will either press it to select the desired item from a list of items through which the computer is scrolling, or use multiple switches to manually ‘move’ through the possible selections and then select the desired item\textsuperscript{7}.

There are few studies measuring actual input or output (speaking) rates of VOCA users, however for text production an output (measured in words per minute - wpm) of less than 5 wpm for switch access\textsuperscript{8}, or less than or around 10 wpm for AAC users using any mode\textsuperscript{9,10} has been suggested. This can be compared to studies of typing on ‘soft’ keyboards that suggest rates of between 8-20wpm\textsuperscript{11}. As the speaking rate of a typical speaker during conversation may be between 100 and 200 wpm\textsuperscript{12,13} it can be seen that current AAC output is in the region of an order of magnitude slower than spoken speech.

People who use AAC and VOCAs often cite challenges and frustrations related to the speed of communication using these systems and also in their simplicity of use\textsuperscript{14,15}. This has been a focus of prior work carried out as part of the NIHR Devices for Dignity (D4D) Healthcare Technology Co-operative\textsuperscript{14,16}. From this and other work, it is clear that an individual’s personal context and environment can be as important as the actual device design on their ability to
communicate effectively. In addition there is an expectation from people who use AAC that devices will become more able to support more responsive, fluent and natural conversations. A range of different approaches have previously been discussed in the literature or implemented in practice as ‘rate enhancement’ techniques. Techniques related to text input include word prediction and completion, abbreviation expansion – a coded input to a phrase bank, and the use of other coding systems such as Morse code. Word completion refers to when a device suggests words based on the input of the initial letters of the word. Word prediction refers to the process when a device actually predicts which word (or words) may follow the current word in a message. Word completion and prediction makes use of a statistical model of language. The model is designed to estimate the distribution of words in a language and is used to predict which word is most likely follow on from the current word or words in a message, given the linguistic context. Garay-Vitoria & Abascal provide a review of text prediction methods.

The selection and output of whole utterances (phrases) has been proposed as a potential method for speeding up communication. However, the use of whole utterances highlights the complexity in the interpretation of the desire for ‘speed’. Although identified by communication aid users as an aspiration, communicative speed is not the only measure of effectiveness in AAC. Words per minute is a throughput measure and is likely to be a poor measure of communicative effectiveness with respect to interpersonal communication and AAC. Light introduces the concept of communicative competence and breaks this into operational, linguistic, social competence and psychosocial factors. Experimental evidence suggests that a quickly produced message with low relevance can reduce the perception of competence in someone using a
McCoy et al. describe the challenge as being to enable “users to quickly and easily retrieve relevant pre-stored messages”.

A wide range of outcomes, from frequency of initiation to quality of life, have been discussed as pertinent to augmentative communication\(^{24,25}\) and although it is beyond the scope of this paper to discuss these, it should be noted that there is not one clear objective related to improving the design of VOCAs. The aims of interpersonal communication also vary greatly and can be characterised variously as: a process of information transmission, as represented by Information Theory\(^{26}\); as self/shared expression involving expression of humour, intimacy or an emotional state\(^{27}\); and as the sharing of personal narratives (story telling)\(^{28,29}\). In addition, Blackstone\(^{30}\) suggests that communication aims, modes and methods can vary greatly between different communication groups or ‘circles’ of communication partners.

1. Context

Communication through spoken conversations has context and this can be conceptualised in a number of ways. A simple model of context would distinguish between external (environmental) context and internal (personal or psychosocial) context. An example of external context may be an individuals’ location or whom they are speaking to; an example of internal context being the intention of the communicative act or the individuals’ emotional state. The context of conversations can effect every aspect of the conversation including the content, use of language, style, tone of voice, intonation, volume, use of slang (colloquialisms) and accent. This effect could be described as the individual’s revealed contextual preferences.

The processing of (revealed) contextual information using algorithms and machine learning techniques is growing within mainstream ICT applications. For example search
engines, such as Google\textsuperscript{31} or iOS 8’s Spotlight\textsuperscript{32} use contextual information, such as where the user is, to inform their search algorithm to influence search results and advertising\textsuperscript{33}. Word prediction engines on smart phones and other devices are also becoming ‘context aware’ – for example, the iOS8 word prediction keyboard claims to take into account “your text messaging style”, “who you’re writing to”, “what the conversation is about” and “how you email”\textsuperscript{34}.

1.1. The Use of Context for Augmentative Communication

People who use AAC may be able to process and use the context of a conversation receptively in the same way as any other individual. However, their operational use of this contextual information is much more limited as their language generation is mediated by use of their communication aid.

An individual using a communication aid may choose to adapt their communication based on the context of a conversation, however this is likely to be in a much more limited way than that of someone using natural speech. An individual using a communication aid that has a fixed set of words (vocabulary) may not have access to the very personalised, low-frequency, or context-specific words required to adapt their communication\textsuperscript{35}. For those using alphabetic typing, non-adaptive word prediction may result in a more static use of language and may poorly predict fringe vocabulary\textsuperscript{36}. Finally, in all cases, constructing a message using a communication aid is slow and this is likely to mean that the utterances are shorter (or telegraphic) and that the message is thus less likely to be adapted to the context of the conversation.

Many of the other expressive changes that context may affect are not available to an individual using a VOCA. Pullin and Cook\textsuperscript{37,38} discuss the challenges of tone of voice when using AAC. In an experiment with 40 participants Pullin and Cook\textsuperscript{38} identified 257 distinct descriptors
of tone of voice and four different perspectives of tone of voice: emotional state; conversational intent; social context; and vocal qualities. Tone of voice within VOCAs is currently limited to either a rising or falling intonation.

Presently there is little incorporation of contextual information in the operation of communication aids. Current aids can be characterised as being equivalent to a physical keyboard - in that they process a sequence of inputs to produce a sequence of words - that can then be displayed or spoken using a speech synthesiser.

Within current VOCA software, the possible approaches to dealing with contextual information can be characterised as:

1. Opening (pages of) pre-programmed content: this is relatively easy to achieve with existing VOCA software architectures. Some examples of the use of location data can be found in commercially available systems – possibly because of the easy availability of location data in mobile computing platforms. For example, Talk Rocket Go™ and Chatable™ use GPS location to load up specific, pre-prepared, vocabulary pages that have been previously linked by the user to that location.

2. Switching topic or vocabulary lists: Some VOCA software include topic or vocabulary lists that can be either pre-programmed or manually stored by the user. These could potentially be loaded based on knowledge of the speaker or other context, yet the current software architectures do not easily support this.

3. Switching language models: Storing and loading alternative language models or ‘user dictionaries’ would be possible within existing VOCA software, though again the current software architectures do not easily support this.
1.2. Incorporating Context Support in Communication Aids

Much of the previous research relating to context support within VOCAs has related to the potential uses of Natural Language Processing (NLP) and Generation (NLG).

The use of adapted language models is an example of NLP technology and of incorporating context into a message generation system. For instance, the claims made for the iOS prediction keyboard (described earlier) are likely facilitated by the use of different language models adapting to the user’s text messaging and email style meaning that the abbreviations and colloquialisms that might be appropriate to use in SMS text messages are not suggested when writing an email.

Adapted language models based on contextual information such as with whom the user is conversing, where the user is, or perhaps even the aim or tone of the conversation offer one way for VOCAs to take advantage of contextual information. As an example, an adaptive system should make different predictions when the user is sitting in a bar with their friends compared to when they are meeting their manager while at work. A model of contextual information for NLP that encompasses location, time, language, communication partner, and (partner’s) conversation content has been presented

NLG is a method for creating natural utterances based on small amounts of input data and is sometimes described as ‘data to text’. These data, much of which could be considered to be contextual data, could be generated by algorithms, gathered from sensors or input by the user. A number of prototype/research NLG systems have been built that use contextual information in varying ways. In some cases this contextual information has been automatically generated and in some it has been input manually by the user. For example, in their prototype system Black et al. used RFID tags to input information on where a child was in a school and swipe cards to
log who they were with; Wisenburn and Higginbotham used speech recognition to attempt to 
recognise the content of communication partners’ speech – i.e. attempting to use a topic as 
context; while Netzer and Elhadad created a system using the Bliss symbolic language as the 
input and used NLG to produce natural sounding spoken sentences.

2. Alternative Approaches to Context Support in Communication Aids

Contextual information can be applied within AAC-mediated interactions in a number of 
ways. We propose the model described in Table 1 where context support is broken down into 
two factors - detection and adaptation:

A. Firstly, the device may (or may not) detect the contextual information 
advertically (e.g. with sensors such as GPS) or it may be manually entered by 
the user;

B. Secondly, the device may (or may not) adapt to the contextual information in a 
number of ways - either by presenting the information to the user to evaluate and 
use, by processing the context in some way and changing the interface to present 
different opportunities to the user, or by processing the context and using natural 
language generation. Looking at Table 1, it is clear that current research has 
concentrated on the use of NLG (modes 2c and 3c). The following section 
provides examples of projects with which the authors have been involved where 
contextual information has been used. These example projects highlight a range of 
different ways of using contextual information to support the use of AAC.

INSERT TABLE 1 HERE
2.1. Example 1: Static context - Intensive Care Units

Context can be situational and fixed (mode 1) – an example of this is a project currently being undertaken related to the use of AAC within Intensive Care Units (ICU). Patients within these units are within a very specific context: they are temporarily unable to communicate because of intubation; are likely to recover speech (when the tube in their airway is removed); have no experience of use of AAC; are in highly stressful environments; and are likely to be in extreme psychological states.

Within this project we are examining the specific contextual factors that influence the use of AAC in this extreme communication environment using a number of methods: a systematic review\textsuperscript{45}; qualitative interview study of ex ICU patients and current ICU staff; and observational data gathered by a designer embedded for a period in an ICU unit. Results to date have suggested a range of specific requirements associated with this context, and this has translated into content requirements - for example in identifying possible topics of conversation to be included in a possible solution - and environmental considerations such as staff training and a design sympathetic to both the ICU patient and staff. Others\textsuperscript{46} have previously demonstrated the challenges of implementing an AAC device in the ICU context.

The outcome of this work will be the design of a new, relatively ‘frugal’ AAC device which will be used within a large scale study to further examine the needs, requirements and barriers to use of these technologies in ICU to inform future technology adoption by the NHS.

2.2. Example 2: Providing Context as Communication Support - VIVOCA

We have carried out research with individuals who have dysarthric speech which is hard or impossible to understand for those who do not know the individual well. This research has
demonstrated an overwhelming desire to rely on natural speech as the primary form of communication - no matter how hard this is for others to understand\textsuperscript{2,47}. This research also suggested that the same desire to use all other methods (such as gesture) before ‘resorting’ to a VOCA was also present in expert communication aid users who had no functional speech.

This reliance upon the use of natural speech leads to frequent breakdowns in conversations. Using conversation analysis, Bloch & Wilkinson\textsuperscript{48} have studied the conversation of dyads, in which one of the participants has dysarthria, and identified sources of breakdown and the challenge in repairing these conversations. As one of the participants in our research said, it can be ‘like a car crash’ once a conversation has broken down and many participants described situations when they had chosen not to attempt to repair the conversation.

Our research, and the work of Bloch and Wilkinson, highlights another potential role for context - that of establishing the key contextual information within a conversation as a way of improving understanding (of the communication partner) and assisting in the repair of conversations.

We have designed an alternative approach to communication which allows the person with dysarthria to provide contextual information, using a VOCA, to support a conversation that is predominately spoken (mode 2a or 2b). The VOCA displays a hierarchy of contextual information that the individual with dysarthria can navigate. As the individual navigates the hierarchy, at each stage the communication aid provides information in the form of spoken phrases which incrementally narrow down the context. For example as part of a conversation which has broken down an individual may navigate through the hierarchy and the VOCA would say: ‘I am talking about a place’ then ‘I am talking about a shop’ and finally ‘I am talking about Mothercare’.
This approach allows the individual with dysarthria to provide key utterances to add the required context into a conversation. It provides many opportunities for conversational repair and supports the communication partner (listener) to understand the context and thus improve their understanding. This approach shifts the onus from the person with dysarthria being ‘at fault’ to giving the communication partner responsibility for repairing the conversation and maintaining their attention on the individual’s speech. This approach will be further explored in future work by the authors.

2.3. Example 3: Adding Specific Contextual Information - SpeakerID

A simple analysis of the external contextual factors relevant in a conversation suggests that the communication partner (speaker) is highly relevant.

We are carrying out a small proof of concept project to attempt to use simple speech technology techniques to identify the communication partner within a conversation. We intend to create prototype ‘SpeakerID’ software that will detect to whom the person using the communication aid is speaking to and to link this into existing VOCA software (modes 3a or 3b).

The SpeakerID software uses speaker-specific information such as the long-term spectrum and fundamental frequency. Each speaker is introduced to the system with a short segment of speech, which is analysed and the characteristics of the speaker are stored for later reference. The software then processes an incoming speech signal to determine to which of the known speakers the signal is most similar too. Proposing a SpeakerID system presents a number of considerations.

Privacy and data protection questions are raised when considering the use of this data: who ‘owns’ the data, how the data is stored (e.g. as recordings, or numerical data), how the data
is processed (e.g. if it is linked to other data such as name, place etc.) and how it is used (e.g. if the VOCA software identifies prior topics of conversation that may be unwanted or negative).

Dealing with errors, and with measures of confidence in the results is also likely to be a key challenge in the integration of this data into a VOCA interface. It is likely, in parallel with other speech technology, that the system would have less than 100% reliability in terms of correct identification. Errors have been identified in other studies as a key issue in the effectiveness of use of speech recognition.

A further question relates to ‘what should, or could, be done with this information?’ A ‘full NLG’ approach of generating a whole utterance based on knowledge of the speaker could be envisaged, nevertheless, without additional contextual information (e.g. place, time, aim or style of the conversation) it is likely these utterances would be perceived as having low competence. This additional information could be added by the user, although several other ‘context-influenced’ approaches may also be considered:

1. Style - changing the style of utterance based on the person or category of person (e.g. if they are known or unknown). For example: known person = informal style - “alright Frank”, unknown person = formal style - “Hello, I am Joe, nice to meet you”.

2. Content - changing content based on person, or category of person.
   a. Personal knowledge: For example: family member or carer - “Can I have a drink”, unknown person - “Can I have a cup of tea, with milk, not too hot, and could you hold it up so I can drink it with a straw”.
   b. Conversational topic: For example: known person - “Nightmare lectures today”, unknown person - “I go to Sheffield University and am studying Design, how about you?”
3. Topic generation - knowledge of past/common topics for each speaker would allow a system to suggest topic dictionaries or vocabularies based on prior use with that person. For example: “Hi James, is your motorbike fixed yet?”

4. Narrative generation - providing access to story-telling, based on previous topics and stories, for example by simply re-offering prior utterances, or stories. For example: “The other week Jack dressed up for charity? They raised lots of money...”

SpeakerID will be used as a critical artefact\(^{50}\) - a physical object that rather than solving a problem aims to stimulate users and industry to examine the potential role and requirements of context support within VOCAs.

3. Conclusions

This paper has attempted to extend the consideration of the potential role of the use of contextual information for support within the design of communication aids. Examples of different approaches to contextual information have been provided: designing a communication aid to deal with the constraints of a specific context; an alternative approach to communication support for individuals with dysarthria where contextual information is presented to the communication partner in order to assist with repair of the (spoken) conversation; and an example of adapting the interface or interaction based on the knowledge of (only the) communication partner.

There are few examples of context support within existing communication aids and the limited research in this area has concentrated on the use of contextual information as an input to natural language generation. We propose that the role of contextual information should be extended and examined further to establish answers to the following research questions, namely:
1. What element(s) of context is (are) most important to consider in relation to improving communication;

2. How contextual information can be used to create more intuitive and simpler interfaces;

3. How context can be used best within communication aids to maximise an individual’s communication (over the most relevant measure for that individual);

4. How communication aid users feel about the use of this information, and,

5. How the possible benefits, risks or disadvantages are perceived by communication aid users.
### Tables

<table>
<thead>
<tr>
<th>Mode</th>
<th>A. Context Detection</th>
<th>B. Adaptation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None (static)</td>
<td>None</td>
<td>VOCA adapted to a static context - i.e. the context is given a-priori</td>
</tr>
<tr>
<td>2</td>
<td>User Driven</td>
<td>User driven context adaptation of VOCA output</td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>Manual / Pre-defined</td>
<td>User inputs the context, the VOCA displays pre-defined content and speaks what the user selects.</td>
<td></td>
</tr>
<tr>
<td>2b</td>
<td>Automatic - using algorithms or NLP</td>
<td>User inputs the context, the VOCA processes and displays this and speaks what the user selects.</td>
<td></td>
</tr>
<tr>
<td>2c</td>
<td>Automatic - using NLP &amp; NLG</td>
<td>User inputs the context, the VOCA processes this using NLP and outputs a natural-language utterance (NLG) from the user’s selections.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Automatic</td>
<td>Context aware: automatic detection of context and adaptation of VOCA output.</td>
<td></td>
</tr>
<tr>
<td>3a</td>
<td>Manual / Pre-defined</td>
<td>VOCA acquires context the VOCA displays pre-defined content and speaks what the user selects.</td>
<td></td>
</tr>
<tr>
<td>3b</td>
<td>Automatic - using algorithms or NLP</td>
<td>VOCA acquires context, processes and displays this and speaks what the user selects.</td>
<td></td>
</tr>
<tr>
<td>3c</td>
<td>Automatic - using NLP &amp; NLG</td>
<td>VOCA acquires context, processes this using NLP and outputs a natural-language utterance (NLG) from the user’s selections.</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. A model of context support in communication aids
4. References


http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6773024


