Speech Driven Environmental Control Systems - A qualitative analysis of the perceptions of users

Simon Judge

Barnsley Assistive Technology Team, Barnsley Hospital, Gawber Road, Barnsley, South Yorkshire, S75 2EP.

(+44)1226 432159

simon.judge@nhs.net

Zoë Robertson

Barnsley Assistive Technology Team, Barnsley Hospital, Gawber Road, Barnsley, South Yorkshire, S75 2EP.

(+44)1226 432159

zoe.robertson@nhs.net

Professor Mark Hawley

Barnsley Hospital & University of Sheffield School of Health and Related Research (ScHARR), Regent Court, Regent Street, Sheffield, S1 4DA.

Mark.Hawley@sheffield.ac.uk

Professor Pam Enderby

Professor of Community Rehabilitation, School of Health and Related Research (ScHARR), University of Sheffield, Regent Court, Regent Street, Sheffield, S1 4DA

Speech Driven Environmental Control Systems – A Qualitative Analysis of users' perceptions

Abstract

Purpose: To explore users' experiences and perceptions of speech driven environmental control systems as part of a larger project aiming to develop a new speech driven environmental control system. The motivation for this part of the project was to add to the evidence base for the use of speech driven environmental control systems and to determine the key design specifications for a new speech driven system from a user's perspective.

Method: Semi-structured interviews were conducted with twelve users of speech driven environmental control systems from around the UK. These interviews were transcribed and analysed using a qualitative method based on framework analysis.

Results: Reliability is the main influence on the use of speech driven environmental control systems. All the participants gave examples of occasions when their speech driven system was unreliable; in some instances this unreliability was reported as not being a problem (e.g. for changing television channels) however it was perceived as a problem for more safety critical functions (e.g. opening a door). Reliability was cited by participants as the reason for using a switch operated system as back up. Benefits of speech driven systems focussed on: speech operation enabling access when other

methods were not possible; quicker operation and better aesthetic considerations. Overall there was a perception of increased independence from the use of speech driven environmental control.

Conclusions: In general speech was considered a useful method of operating environmental controls by the participants interviewed however their perceptions regarding reliability often influenced their decision to have backup or alternative systems for certain functions.

Introduction

Environmental control systems were initially developed in the 1960's [1] for patients who had suffered spinal cord injury. The aim of these systems and environmental control systems today is to enable people with disabilities to operate equipment within their immediate environment, enabling them to do things such as answer an intercom, open the door, call for help or operate the television etc.

The first systems were large wall mounted devices connected by cables to the devices they were operating (hard-wired). Significant developments in environmental control have occurred since the 60s [1][2][3] predominantly driven by advances in technology such as radio and infrared transceivers, transistors and micro-controllers. These advances have enabled devices to become smaller and portable, perform more functions and no longer be hard-wired to the peripheral devices.

Despite these advances, the main methods of operation of environmental control systems have changed very little; the majority of systems are operated either via a single switch, a number of switches or via a keypad on the device. Whilst keypad input can be an efficient method of control for those whose physical ability is relatively good, this is not the case for many people who have need of environmental control systems. The majority of environmental control systems are controlled by switch access which can be slow, effortful and frustrating. Speech is a possible control method for many current and potential users of environmental controls or other Assistive Technology [4] and can make environmental control faster, less effortful and less frustrating [5]. Existing speech driven environmental controls appear to have a very low prescription rate however [6]. Speech control of personal computers is more widely known and used within the field [7][8], suggesting that speech control is acceptable to users. The question of why speech is not widely used as a control method for environmental control systems remains open.

The reasons for carrying out this study were to explore the question of why speech is not used more widely as a control method for environmental controls and further the knowledge in this poorly evidenced area of Assistive Technology. Anecdotal evidence exists for the lack of prescription and adoption of speech-driven systems however no studies have looked at user's perceptions and opinions on speech as an access method for environmental controls. In addition it was planned that results from this study would provide user input to the SPECS project [9]. The SPECS project is a large scale

collaborative project with the aim of developing a new speech driven environmental control.

This paper presents the results of a qualitative analysis of semi-structured interviews with participants currently using speech driven environmental controllers.

Method

Recruitment

United Kingdom ethical approval was granted by North Sheffield Ethics Office. Participants for the study were identified by contacting health professionals involved in the prescription of environmental control systems via an assistive technology professionals' mailing list (www.jiscmail.ac.uk/lists/assistech.html). Professionals were sought who could provide details of clients in their areas currently using speech driven environmental control systems, or clients who had tried out speech driven environmental control systems but had rejected them. Eleven professionals expressed an interest and contact was pursued with five for participant recruitment. As well as providing participants for the study, data were also collected from these professionals by way of interviews and focus groups. Data from professionals were collected to help understand the prescription rationale for speech driven environmental control systems; this data is to be presented in a separate paper.

The inclusion criteria for this opportunity sample of participants were that they should be over 16 years of age and either be people who currently use

speech-driven environmental controls or people who had tried speech-driven environmental controls and rejected them. In addition participants needed to have sufficient cognitive ability to take part in the interview, i.e. to understand the questions and prompts set out in the topic guide.

Interview procedure

Interviews were carried out at the participants' homes by researchers experienced in Assistive Technology and trained in qualitative interview techniques. The interviews lasted approximately an hour in length and were designed to be open and free ranging whilst also drawing on a pre-defined topic guide. The topic guide was developed following a literature review and discussion between members of a small expert group. It was designed to capture participants' experiences and issues with existing speech driven environmental control, including the features they like, dislike or would like. Each participant was interviewed once. Interviews were recorded using a digital voice recorder and transcribed by an independent transcription service.

Initially two speech driven environmental control users were interviewed. Data from these initial interviews were analysed using a framework analysis approach, as detailed in the data analysis section. This initial analysis identified areas to be explored in further interviews and ensured that the topic guide was appropriate to elicit the type of data required. No changes were deemed to be necessary to the topic guide following this initial review and so the rest of the user interviews were completed and analysed.

Data Analysis

Framework analysis [10], a qualitative research approach, was chosen as the basis for interpretation of this user data since it allows a focused analysis. A framework comprises main themes and sub-themes of common perceptions. Main themes were determined by the needs of the overarching project to develop a new speech driven environmental control system; sub-themes were drawn out from the data over the stages described below. As an example, a main theme is 'Factors influencing failure' and a sub-theme is 'Changes in voice'.

The main themes, developed from the project and research specification, were defined as:

- Factors influencing failure – negative perceptions of speech driven environmental control
- Factors influencing success – positive perceptions of speech driven environmental control
- Interface – perspectives on interaction with the device
- Usage – perspectives on participants' day-to-day use of the device
- Background – how participants view their general situation

The analysis process was performed with a number of stages as follows:

Initial Framework Development:

- Individual researchers separately identified sub-themes using the data from the initial two interviews.

- Researchers then jointly compared and consolidated the sub-themes into an agreed initial framework. Consensus was agreed by joint discussion between both researchers of each of the sub-themes' definitions with the aim of identifying any commonalities. In many cases sub-themes identified were similar however each researcher had named them slightly differently.

Initial Data Coding:

- The initial two interviews were then coded (extracts were identified in the data which related to each sub-theme) by both researchers independently.
- This coding was then jointly compared and consolidated. The researchers reviewed all of the coded extracts and identified both extracts which had been coded differently by individual researchers and those that had only been coded by one researcher. These extracts were discussed in detail, referring to the definitions of the sub-themes, until consensus was achieved on the correct sub-theme for the extract.
- The framework and topic guide were checked during this process to ensure good coverage of data (extracts) across sub-themes.

Full Data Coding:

- Further interviews were coded following the same procedure i.e. independent coding by each researcher and then joint comparison and consolidation.

- The data in the sub-themes was further reviewed by both researchers and weak themes without coverage across participants were consolidated.
- Summary data (reported in this paper) were then produced by agreeing on representative extracts from the data.

Results

The initial plan was to interview both successful users of speech driven environmental control and those who had tried speech driven environmental control and found it unsuccessful. After contact with professionals it became evident that it would not be possible to recruit users who had rejected speech driven environmental control systems. Reasons for this included: the very small cohort of speech driven environmental control users in different areas; self selection in the prescription of these devices; rejection due to cognitive changes which would also impede them being interviewed about the system; users having died since prescription.

Twelve existing users were recruited, the number being determined by saturation of the data. Table 1 gives an overview of the twelve participants.

A large amount of data (extracts from the interviews) were generated from the data analysis – with each main theme having from 7 to 15 sub-themes and a fair coverage of data across participants.

Factors influencing failure

The main overarching feeling among participants was of a lack of reliability particularly associated with errors in recognition. This can be observed with a finer grain in a number of the sub-themes although the main reason appeared to be due to sound interference.

> **8 -** "normally it doesn't let me down but it did on that occasion when I really needed it, that's the trouble, when I really needed it."

> **3 -** "the trouble also with voice is that if you turn on a CD player and it's playing too loud then you have problems, answering the phone's always a nightmare because you try to say answer phone but the phone's ringing at the same time"

Speech specific reliability issues included: changes in the user's voice affecting the system; the phonetic similarity of words; sound interference; specific characteristics required for user's voices and variability in user's voices and the acoustics due to the environment. A complex web of possible reliability issues was identified, for example: changes in a person's voice over a day or over a longer period of time would cause the device to become less reliable, but in addition so might a change in room or environment:

> **12 -** "So if I haven't had a drink then my voice is that dry that the voice activator doesn't recognise it and you can sometimes scream at it and it will get you nowhere."

Words that are phonetically similar often required modification of words to improve reliability but also participants felt that they had to talk to the device in a certain manner:

7 - "You've got to say 'reda' instead of 'red' or 'greena' and 'bluey' and it just – people in the room are laughing and they're thinking 'what the hell's he saying bluey for?"

6 - "you've got to speak really, really clearly and think before you speak."

Several technical issues arose: difficulty interfacing with some peripherals; difficulty of training new functions; limited menus; problems with batteries & infrared. These issues could be considered as general environmental control issues, although there were some peculiarities specific to speech enabled environmental control.

Human issues were also addressed: lack of patience with system; lack of training, and learning the operation of the device. The general perception was that the systems and support could be frustrating, and this could be detrimental to their potential usefulness.

5 - "You can get frustrated – I don't know what other people are like – I know I get very frustrated with it".

Factors influencing success

Participants for this study were biased, since by their nature they were deemed to be 'successful' users of the systems (since they were still using

them). It is thus interesting that the 'Factors influencing failure' theme was much more heavily referenced – i.e. participants had more negative things to say than positive. Despite this, participants were generally very strong advocates for speech as a method of interaction and this shows two of the qualities highlighted for successful use namely determination and resilience.

10 - "It's a good piece of equipment, it's one of the best bits of equipment I've ever seen."

A number of areas were identified as reasons that participants perceived the device to be positive: simplicity of use; speed of operation; use of voice as an access method when they could not access another system. These sub-themes confirm that speed and simplicity are perceived to be a key feature of speech control. However the data also suggests that currently speech is prescribed as a last resort and often in conjunction with switch accessible systems. Participants perceived the benefits from speech control when they could not (for physical or other reasons) access their alternative environmental control system and also when used as a complement to their other system.

5 - "if anybody is like me I would recommend a combination of the two. It is like stopping and rewinding video and stuff like that, it is quicker to do it by voice"

8 - "it all seemed very daunting at the time, but it was OK - once I picked it up it was not too difficult."

12 - "I suppose now my arm movements have got better but my arms still stiffen up, at night time especially, so that's when the voice comes into operation and I rely on my voice more."

Other areas identified related to ways in which participants had successfully learnt to use the device. These centred around personalising the device and learning compensation tactics: ability to use any word for commands; system training; tailoring the device to the user's needs; using particular voice intonation and patterning (consistency).

6 - "I've got a lot of electrical commands to put in. You put your TV in to get it going on and off, you're talking ten minutes if you train it properly. I've got it down to a fine art now so I can do it really quick."

3 - "it's more sort of TV on, off, channel up, down, words as opposed to yeah. TV would you mind turning on"

Interface

The interface of existing speech devices is restricted to a small text display and limited auditory feedback. Thus, this limited experience may influence their perceptions and this should be noted when considering the participants' perceptions recorded below. Interesting observations were made predominately around the aesthetics of the device, the microphone and also around the feedback and interaction with the device.

Participants generally relied on the auditory feedback of the device and rarely used the display however they could see a reason for a display when this was discussed. Participants perceived that speech was a good interaction method when 'mastered' and that scanning, the main alternative, was slow and tedious.

9 - "I never use the display, I always listen to it speaking back. I don't even bother sometimes with that, I look and see what it's done to the screen. But yes, sometimes it's useful to listen to the speaking back because you know where the errors have gone, you hear the errors."

10 - "A typical woman. If you speak calm, get what you want in the end, shout at them, go off in a huff."

10 - "you have to think where things were and it goes beep, beep, beep stop that then, and then beep, beep, beep and if you want to make a phone call, oh God they were terrible. Beep, beep, beep and you'd have to find the number and you look for the number and you stop it then you go somewhere and it goes sideways and I think 'argh!'"

Participants identified that aesthetics can be an important aspect and noted the disabling image of many Assistive Technology products and potentially very positive and enabling image of speech controlled devices.

3 - "by definition probably in the 16 to 24 age bracket you've probably got another 40, 50 years to tick along and you don't suddenly want to be thrust into some you know, and you don't really want your home to look over you like a hospital ward, you know,

particularly, and so I think the aesthetics also has to play something as to whether people feel comfortable using it"

Speech Driven Environmental Control (SPECS) Usage

The biggest overarching sub-themes related to risk, security and independence – participants overwhelmingly associated the device with providing independence. In addition there was also a process of risk assessment: choosing the appliances that the device controlled and ensuring that human and other backup methods existed. The risk assessment aspect also highlighted that participants appeared to accept a certain level of recognition error, given that it was not going to cause any damage (just maybe turn a TV channel over for example) – there was a trade off between the reliability, risk and functionality of the systems.

12 - "So I've got back-up for both because I do need it, because obviously my voice alters so much, especially in the mornings."

3 - "whereas with the TV the worst thing that could happen is you end up watching the wrong channel or it gets too loud and when somebody then does come to assist you, you haven't threatened your existence."

10 - "The nurses came in the morning, I couldn't let them in and I don't know what had gone wrong with it, but it just went off and there's keys across the road, so they just got the keys and came in."

A second commonality in the sub-themes is around the outcomes of using these environmental control systems – reducing the load on carers and providing independence to participants. These factors strongly support the reasoning behind environmental control prescription in general.

8 - "I can open and shut a curtain, if I want to look at the moon I can do, if I want some fresh air in the room I can open the window, I can put my heater on if I get a bit cold, so it has made a big difference to my role in the house on my own."

1 - "I mean it's not just frustrating for me not being able to do it, it gets frustrating for my husband, when he's in the middle of cooking a meal and I'm saying "P can you come and put me this on, can you come and do my back" anything like that that saves him, you know, is brilliant."

Background

This theme highlighted the context in which existing speech enabled devices are used and highlighted some of the selection bias in prescription towards more able, well supported, clients. The prescribers' perspective on this was captured in data from professionals and this is presented in a further paper.

In general participants viewed themselves as very cognitively able and viewed themselves as being computer literate with eight having previous or current use of voice recognition software on computers.

1 - "I'm a really busy person, I run Avon, I'm an Avon representative and I'm one of the best sellers of things like this, so that's a little kind of thing I do on the side which keeps

me on the go, keeps my mind on the go, I mean I like to do crossword puzzles and things like that."

7 - "My background is in computing so consequently I'm used to things like this"

Ten of the participants had a relatively long history of environmental control use and some used a second switch-operated controller in conjunction with their voice system. All participants had support from carers and generally had a good relationship with their environmental control service. Support ranged from carers who were close relatives and lived with the participant to paid carers who changed frequently. The inability to rely on carers to carry out technical tasks was a common theme.

10 - "This is partly my desire not to force my wife to have to do it because she's very busy doing all sorts of things. I mean not only is she my 24 hour carer but she has various other things to do so it's a matter of – I don't think I've got the right to demand her time."

4 - "Well, I have two environmental control systems. One, which is on this chair here which is not by voice and then one which is by voice."

Discussion

This study has provided an in-depth analysis of the experiences of existing users of speech driven environmental control systems. The study followed a qualitative methodology and the analysis process may be considered

particularly rigorous since at each stage two researchers independently coded and then compared and consolidated the data.

A weakness of the study was that all of the users were ‘successful’ users i.e. they had not abandoned use of the speech driven environmental control system despite any issues with it. Participants still provided both positive and negative views on the device – however they should be considered a biased sample because of their continued use. The data reported on would have been strengthened by collection of data from users who had tried speech operation but rejected it, however as detailed in the methodology it was not possible to do this.

In this study we tried to interview a broad spectrum of users of speech driven environmental controls by identifying users in different geographical areas. All of the participants however were recruited via the same channel (i.e. via NHS services) and this may have also been a source of selection bias. This potential selection bias may have caused data saturation to be reached with lower participant numbers than if more varied channels had been used to source participants.

We have shown that reliability is the key issue when it comes to using speech as an access method – much evidence was found as to the perceived unreliability of the devices currently used by participants. The range of reasons identified illustrates the complexity of the problem and the technical challenge that needs to be overcome in designing an effective method of

voice control. It is interesting to note that reliability featured so strongly, since the selection bias of the participants would suggest that they would all be expert-users and have a low recognition error. In fact, the large amount of data under the 'reasons for failure' theme, despite this selection bias, may be illustrative of why these devices are not more widespread – even 'successful' users struggle with them!

The reasons that participants use existing voice systems at all is seen in a number of areas: firstly the participants place a high value on their use of environmental control and the independence it provides and so persevere with the devices; secondly they are willing to accept a certain level of unreliability as long as it does not put them at risk; thirdly, in many cases, the potential gains from using voice outweigh the annoyance of errors – either because of speed gains or because they were otherwise unable to access conventional systems.

Although these participants all continue to use speech access devices, the experience of a lack of reliability for these non-critical functions influenced their trust in using the devices to control functions associated with more risk, for example making a phone call in an emergency situation, and often led them to have some form of backup for this function.

As mentioned in the introduction there is limited research in the field of speech driven environmental controls to compare this study against. Geggie et al [6] provide anecdotal evidence on their experience of provision of speech

driven environmental controls within their service and cite reliability problems and use with safety-critical functions as important. These experiences concur with some of the issues highlighted by the participants in our study.

A number of themes arose that could be considered to be relevant to general environmental control use, for example, the participants' perceptions of the outcomes of provision. The Matching Person and Technology Model [11] identifies the importance of considering not only what technology can do but the environment in which it will be used and the personality of the user when assessing for assistive technology. Failure to consider factors such as service provision have been documented as significant in the adoption or failure to adopt AT [12], [13]. Studies have also highlighted the lack of user involvement in the design of assistive devices [14]. The background of this study is to develop a design specification for a new speech driven environmental control and hence the analysis of the data collected has focussed on users' perceptions of the technology (speech driven environmental control systems). Data from interviews with professionals will be analysed and published separately.

The data collected in this study demonstrates that speech is a valuable access method for environmental control but suggests reasons why speech driven systems have a low prescription level. Whether the current level of prescription is appropriate or whether professionals are disproportionately put off from trialling these systems remains an open question and the subject of future research.

In addition to providing evidence around the current use of speech controlled environmental control devices, this data will be used to influence the design of a new device. Noting the number of different reasons for possible recognition errors described in the results it will be important to ensure the new device has increased reliability in 'real' situations. Many of the other themes identified relate to specific usability issues which will be considered in the device design with the aim of allowing speech to become an intuitive interaction method.

Clinical Messages

- Successful users of current speech driven environmental control appear to be relatively cognitively able, computer literate and tolerant.
- Successful use of existing systems depends on accepting a certain level of error, use in conjunction with other backup methods and a strong motivation to increase speed or access to the equipment being controlled.
- Improving reliability is key to design of a new speech driven device.

Acknowledgements

This work was undertaken as part of research funded by the National Institute of Health Research (NIHR) Health Technology Devices programme. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health or NIHR.

References

[1] Shaw H. `'Possum': an electronic aid to fuller living for the severely disabled.'. Physiotherapy 1962; 48:308-309.

[2] Vincent C, Drouin G, Routhier F. Examination of new environmental control applications. Assist Technol. 2002;14(2):98-111.

[3] Wellings DJ, Unsworth J. Fortnightly review: Environmental control systems for people with a disability: an update. BMJ 1997; August 315(7105):409-412.

[4] Hawley MS. Speech Recognition as an Input to Electronic Assistive Technology. British Journal of Occupational Therapy 2002: 65(1):15-20.

[5] Hawley, M.S., Enderby, P., Green, P., Cunningham, S.P., Brownsell, S. Carmichael, J., Parker, M., Hatzis, A., O'Neill, P., and Palmer, R. A speech-controlled environmental control system for people with severe dysarthria. Medical Engineering and Physics 2007; 29(5):586–593.

[6] Geggie C. Voice Control of Environmental Control Systems. ACNR 2003; September 3(4).

[7] Sears A, Karat CM, Oseitutu K, Karimullah A, Feng J. Productivity, satisfaction, and interaction strategies of individuals with spinal cord injuries and traditional users interacting with speech recognition software. Universal Access in the Information Society 2001; June 1(1):4-15.

[8] Kotze P, Eloff M, Adesina-Ojo A, Eloff J. Accessible Computer Interaction for People with Disabilities. The case of quadaplegics. ICEIS 2004 – Human-Computer Interaction.

[9] Hawley M, Cunningham S, Judge S, Kolluru B, Robertson Z. Using qualitative research methods to inform user centred design of an innovative Assistive Technology device. Proceedings of CWUAAT 2008. In press.

[10] Ritchie J, Lewis J. Qualitative Research Practice: A guide for Social Science Students and Researchers. SAGE Publications Ltd. 2006.

[11] Scherer MJ, Craddock G. Matching Person & Technology (MPT) assessment process. Technology and Disability. 2002 January;14(3):125-131.

[12] Scherer M, Sax C, Vanbiervliet A, Cushman L, Scherer J. Predictors of assistive technology use: The importance of personal and psychosocial factors. Disability and Rehabilitation. 2005;27(21):1321-1331.

[13] Phillips B, Zhao H. Predictors of assistive technology abandonment. Assist Technol. 1993;5(1):36-45.

[14] Salminen AL, Petrie H, Ryan S. Impact of computer augmented communication on the daily lives of speech-impaired children. Part I: Daily communication and activities.

Table 1 – Participant details

Participant Code	Age	Diagnosis	Length of time using speech driven environmental controls	Main environmental control access method	Other environmental control experience
1	68	Motor Neuron Disease	Approx 6 months	Voice	Previous experience of switch operation
2	51	Spinal Cord Injury	Approx 2 years	Voice and Switch	None
3		Spinal Cord Injury	Approx 7 years	Voice	Scanning device for backup and previous experience of switch controlled system
4		Multiple Sclerosis	Approx 2 years	Voice and Switch	None
5	47	Spinal Cord Injury	Approx 3 years	Voice and Switch	None
6	36	Spinal Cord Injury	Approx 4 years	Voice and Direct Access	None
7	42	Arnold-Chiari Syndrome	Approx 2 years	Voice	Previous experience of switch controlled system
8	56	Spinal Cord Injury	Approx 2 years	Voice and Direct Access	None
9		Motor Neurone Disease	Approx 2 years	Voice	None
10	55	Spinal Cord Injury	Approx 3 years	Voice	Previous experience of switch controlled system
11		Quadraplegic	Approx 2 weeks	Switch	Trial of voice controlled system
12	45	Spinal Cord Injury	Approx 5 years	Voice, Switch and Direct Access	None