promoting access to White Rose research papers



Universities of Leeds, Sheffield and York http://eprints.whiterose.ac.uk/

This is an author produced version of a paper published in **Journal of Assistive Technologies**.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/74407

Published paper

Judge, S., Robertson, Z., Hawley, M. (2011) *The limitations of speech control:* perceptions of provision of speech-driven environmental controls, Journal of Assistive Technologies, 5 (1), pp. 4-11 http://dx.doi.org/10.5042/jat.2011.0096

Authors

Simon Judge

Senior Clinical Scientist, Barnsley Assistive Technology Team, Barnsley Hospital NHS Foundation Trust, Gawber Road, Barnsley, S75 2EP. UK simon.judge@nhs.net

2007 to date: Senior Clinical Scientist, Barnsley Assistive Technology Team. 2004-2007: Clinical Scientist, Access to Communication and Technology, Birmingham MEng Electronic and Electrical Engineering State Registered Clinical Scientist

Zoë Robertson

Assistive Technology Clinical Specialist, Barnsley Assistive Technology Team, Barnsley Hospital NHS Foundation Trust, Barnsley, S75 2EP. UK zoe.robertson@nhs.net

2007 to date: Assistive Technology Clinical Specialist, Barnsley Assistive Technology Team 2001-2007: Bioengineer, Gait Analysis Laboratory, Derbyshire Royal Infirmary

> MEng Integrated Engineering State Registered Clinical Scientist

Mark S. Hawley

Professor of Health Services Research, School of Health and Related Research, University of Sheffield, Sheffield S1 4DA, UK & Barnsley Hospital NHS Foundation Trust

Mark.Hawley@sheffield.ac.uk

2007 onwards: Professor of Health Services Research, University of Sheffield 2000 onwards: Head of Medical Physics and Clinical Engineering, Barnsley Hospital

> BSc, PhD, CSci, MIPEM, Hon. FRCSLT State Registered Clinical Scientist

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.

The Limitations of Speech Control: Assistive Technology Professionals' Perceptions of Speech-Driven Environmental Control Systems

Abstract

Aim: This study set out to collect data from assistive technology professionals about their provision of speech-driven environmental control systems. This study is part of a larger study looking at developing a new speech-driven environmental control system.

Method: A focus group for assistive technology professionals was conducted. This focus group was recorded, transcribed and then analysed using a framework approach.

Results: The analysis suggested that professionals have a 'mental model' of a successful user of a speech-driven system and that in general they consider such systems either as a 'last resort' or to work in parallel with another system as a backup. Perceived poor reliability was highlighted as a major influence in the provision of speech-driven environmental control systems although there were also positive perceptions about the use of speech under controlled circumstances. Comparison with published data from end-users showed that professionals highlighted the majority of issues identified by end-users.

Conclusion: Assistive technology professionals think that speech has potential as an access method but are cautious about using speech-driven environmental control systems predominantly due to concerns about reliability. Professionals seem able to empathise well with the challenges faced by end-users in use of these systems.

Keywords: Environmental Control Systems, Speech Recognition, Provision, Perceptions, Electronic Assistive Technology

Introduction

Assistive technology enables people with disabilities to achieve tasks that they may otherwise be unable to do. Environmental control systems are one of the most well established forms of electronic assistive technology in the United Kingdom. They provide people with severe disabilities the ability to control their home environment, for example operating the television or answering the phone. In the United Kingdom, environmental control systems are generally provided through NHS electronic assistive technology services or community occupational therapy teams. Generally assessments for provision of systems are carried out by clinical scientists or specialist occupational therapists and provided to people with severe disabilities affecting the use of their upper limbs. Due to the nature of the target user's disabilities such systems are generally controlled (accessed) using a single switch – this can potentially provide access to the full range of household equipment using even very small functional movements.

A small amount of previous work has looked at the benefits and improvements to the quality of life for individuals provided with environmental control systems, for example Palmer & Seale (2007) used grounded theory to analyse data from interviews with system users. A review of the literature by Craig et al. (2005) however, noted the lack of systematic study in this area. A number of other small studies have looked at issues around service delivery, for example Holme et al. (1997) and Williams & Bowie (1993), both identify

audits of provision levels, for example Paul et al. (2006), Maguire et al. (2001),

Novak (1998) and McDonald & St Leger (1996).

Significant technological advances have occurred within the control units as processing power has developed, as reported in Wellings & Unsworth (1997) and Vincent et al. (2002). However there has been little progression in the method of control of these systems since their advent in the 1960s (Shaw 1962). Anecdotal evidence suggests that the current predominate method of control of environmental control systems is "single switch auto scanning".

Further description of the range of different possible access methods can be found in Judge & Colven (2006). With "single switch auto scanning" the device displays a list of options, the user activates a switch and the device starts moving one at a time through the list of options (scanning). When the desired option is reached the person activates the switch again to make a selection.

This method of control can be slow and laborious, as identified in Judge et al. (2009), and may contribute to the rate of abandonment of environmental control systems.

Speech control is an obvious alternative to existing methods of control, offering a potentially intuitive, quicker and more effective method of control. Geggie (2003) and other anecdotal evidence suggests that there is one main speech-driven controller in use in the UK and that this has a low rate of provision.

This work was carried out as part of a larger project called SPECS (Hawley et al. 2007) which is developing a new generation of speech-driven assistive devices. The objective of this stage of the project was to develop a specification for a new device via consultation with current end-users of speech-driven environmental control systems and professionals involved with the provision of these devices. This paper reports the data gathered from professionals at a focus group for the project. The results of the end-user consultations are reported in Judge et al. (2009).

Method

Recruitment

Assistive technology professionals were defined as anyone who would recommend or prescribe an environmental control system as part of their day to day role or had significant professional involvement with end-users of speech-driven systems. Professionals were recruited to the study through the 'Assistech' mailing list, an open email mailing list subscribed to by over 200 UK assistive technology professionals which is independently hosted through the academic JISCMAIL network (https://www.jiscmail.ac.uk/cgi-bin/webadmin?SUBED1=assistech&A=1). A request for participation was sent to the Assistech list and a number of professionals replied. Inclusion criteria were that professionals had current or past clients that used speech-driven systems.

National ethical approval for this work was obtained from North Sheffield Research Ethics Committee.

Procedure

A topic guide was developed for use at the professionals' focus group and two exploratory interviews were carried out in order to test the topic guide's validity. This topic guide was then used to structure the focus group session. A focus group was chosen as the method for collecting this data in order to allow professionals to share and explore each other's experiences in the use of speech-driven environmental control systems. The aim of using a focus group was to provide a richer data source in a field where collaborative work or discussions could be unlikely due to the low incidence of provision. This rationale is supported in the literature e.g. Ritchie & Lewis (2003).

Six professionals attended the focus group which was held in Sheffield and included National Health Service clinical scientists, clinical technologists and educational staff [see Table 1] with an average of 13 years' experience in electronic assistive technology. The focus group was split into two parts over the course of a day: the morning session explored the participants' issues with existing speech-driven environmental control systems and the afternoon session explored the participants' aspirations for speech-driven environmental control systems. The morning session concentrated on existing devices and the main headings in the topic guide were: What are your experiences of speech driven environmental controls; What are the advantages of current speech driven environmental controls; What are the disadvantages of current speech driven environmental controls; How do current devices compare with other voice controlled applications e.g. computer access? In your experience what influences the successful or unsuccessful experience of a user of a speech driven environmental control; Why do EC professionals tend to be

biased against speech driven environmental controls? The afternoon session looked at aspirations and the main headings in the topic guide were: If you could start with a blank page what would be the features you would design into a new EC device; What factors of existing speech driven environmental controls and switch devices would you incorporate into a new device; What do you feel are the most important features of a speech driven environmental control?

The focus group was facilitated by an experienced researcher with good knowledge of assistive technology but who was independent of the project. Participants were not asked to identify themselves before speaking as it was felt that this may be detrimental to free-flowing conversation between participants.

Data Analysis

The focus group was recorded, independently transcribed and analysed using a qualitative method based on Framework Analysis (Ritchie & Lewis 2003). Using this method, the transcriptions are analysed and excerpts identified that represent certain perceptions or views. Excerpts are then 'coded' into a framework of main themes and sub-themes.

A framework developed through analysis of end-user data was chosen since it had been developed with the same aims and objectives in mind but with different participants (end-users). It was also chosen to allow comparison between the perceptions of the end-users and professionals. Researchers also considered the creation of new sub-themes if existing sub-themes were

not representative of some data. The data were independently coded by the two researchers and then collaboratively merged and consolidated. The Framework Analysis methodology used is described in greater depth in Judge et al. (2009).

Results

This paper concentrates on the topics most strongly referenced by the participants as being relevant when assessing for, and considering provision of, speech-driven environmental control systems. Five main themes were developed in the analysis and these are presented here with the sub-themes illustrating the issues in depth. The main themes in the framework are: background; speech-driven environmental control usage; interface; factors influencing success; and factors influencing failure. Sub-themes are *italicised* when discussed, extracts from the data are shaded and names have been replaced by X to preserve anonymity.

'Background' Theme

The background and situation of an end-user was strongly considered by participants when considering someone's likely success with a speech-driven environmental control system. 'Disability or Condition' and 'Cognitive Ability' were both strongly referenced. Good memory and understanding were the main requirements cited by participants. These requirements were reflected in the perceptions of which end-users were most commonly provided with systems: people who had had a spinal cord injury were highlighted as likely to be provided with a system whereas end-users with a condition with cognitive involvement were considered unlikely to be suitable for provision.

I looked at the patients who I remembered had got them and they were all spinal chord injuries

You've got to match the client's cognitive profile haven't you and obviously it's not always appropriate.

The low level of provision of speech-driven environmental control systems within services was also noted and the difficulties with systems, described elsewhere in this paper, were agreed as the reasons for this. Participants only had experience prescribing one type of system, tallying with the experience noted in Geggie (2003) indicating that this is the main system currently available in the UK. The level of support required for speech-driven environmental control systems was also discussed with the implication being that only services with the ability to provide a high level of support would consider installing them.

Obviously most of us work in services that have in-house technicians so the type of support that X's talking about where you may have an iterative process to get somebody up to full speed, we can actually provide that support.

'Speech-Driven Environmental Control Usage' Theme

Participants strongly identified case scenarios when they might use speechdriven environmental control systems. Two predominant situations emerged:

- Use as a last resort 'used when switch input is not acceptable', when end-users had been unable to successfully use an alternative more conventional access method.
- Use as a 'backup device' either in conjunction with a switch system or for times when the switch system could not be used.

she spasmed at night and she moved away from the switch and she got fed up with this so she just didn't use anything. So after about a year I introduced the speech recognition again

we've got eleven now and it's literally just down to that issue of when there is no other available controllable function

he was turned every two hours as well because of pressure sores, so he didn't want to go from the switch. We initially set him up with two switches but the nursing staff didn't have the patience to actually reconnect the system, so he went for a voice

The issue of risk assessment was also raised by participants who felt that the use of speech-driven environmental control systems needed to be carefully considered in terms of 'risk assessment of functions to control'. This subtheme is closely related to the perceptions of the participants that speech control is a last resort consideration and that the systems are unreliable. Participants did not rule out installation of speech-driven environmental control systems on the basis of risk, but noted that they carefully consider the situation, possible alternative systems and try to ensure backup systems are in place or installed alongside. The use of 'backup devices' in conjunction with speech-driven systems further demonstrated the participants' reluctance to install a speech-driven environmental control system as an end-user's only environmental control system.

So again, we're talking about risk analysis here and I'm very reluctant to put in a speech recognition system without some form of control back-up.

Well the two people I've got at the moment that use the critical function, the alarm, they've never come back to me as it being a problem.

You can make the risk assessment argument that you haven't got any other choice.

Speech-driven environmental control systems were described as needing 'extensive training' to set up, indicating some of the additional demands on services when providing these devices.

It's actually a more demanding process than actually using the device for people who have got any respiratory restrictions, so it's a good indicator to final use if they're able to train it.

for somebody who understands the process and has done it once before, you can do a pretty complex one in half an hour without too many problems. Maybe thirty or forty words to train, that sort of thing.

The usage of the device was also linked to the end-user's abilities when participants referred to the 'menu structure' of the device. This menu structure determines how the device is used by the end-user and participants felt it added a certain cognitive load to the devices' use.

I mean the cognition issue is definitely through the mapping process and controlling it.

It all depends on the way it's set up and the structure and the way that you link through the menus so that's all down to interface with the client isn't it?

'Interface' Theme

The control method of speech-driven environmental control systems was covered by the participants. The spoken input to these devices was

discussed and it was strongly felt that participants would screen possible endusers for particular characteristics of their speech and voices. In particular it was felt that end-users would need to learn to *'understand how to talk to the device'*.

Consistency is certainly key and as I say, I've detected characteristics in two of the devices that make it perform better.

Participants showed their appreciation of end-users' requirements by drawing attention to the importance of the 'aesthetics' of the currently available device (which is relatively small and discrete).

there's an image aspect that the market has created and people do latch onto that I'm sure, but I think vanity one is there as well. You do get a clean system in a lot of cases. No intruding switch.

Day to day practicalities of controlling the device were covered with the importance and potential problems of 'microphones' being highlighted.

Unless you have to wear a head microphone which comes with its own overhead doesn't it?

You've suddenly got to put it on but if you've got a spasm of your head it's not going to reposition itself.

'Factors influencing Success' Theme

Participants were generally positive about the use of speech-driven environmental control systems when used in certain carefully controlled situations. A number of factors emerged that participants perceived as positive indicators of use of speech-driven environmental control systems. 'Using particular voice patterning and consistency' and 'Modifying control'

words to be phonetically distinct' related to techniques that the participants had learnt to maximise the chance of success for end-users.

I think it likes hard pronounced syllables and that's what it performs best on, hence when you command it you emphasise the hard bits on each word.

Yes, it's experience really. You've got to look at extending the word or having a different sounding at each end of the word. We use 'alarm alarm' to do 'alarm' because it wouldn't accept 'alarm' on its own. So you've got to pick up tricks that will work.

Participants confirmed the situations when speech-driven environmental control systems can be successful i.e. the 'use of voice when unable to access other system'.

The simplicity of speech as an interaction method was also discussed in a positive light and suggested as a reason why speech can be successful. In addition, the ability to *'use any word for commands'* was viewed in a positive light by participants.

'Factors Influencing Failure' Theme

'Reliability' was agreed by participants to be the main reason that speechdriven environmental control systems were likely to fail.

Then it's a question of recognition accuracy. If devices recognise it and it's accurate and it does the job then it's all right but my small experience of using a commercial environmental control system, which is only just the one device, it just couldn't recognise my voice so I had not much confidence in it whatsoever, let alone giving it to somebody who's got a disability.

'Sound Interference' was also discussed and referred to as a source of reliability problems, including interference from TVs, telephones and other people.

I think that's a base issue which has got to be got over because if the television on, microphones tend to pick up these extraneous noises

Participants also identified some factors related to end-users that led to failure, including: the rigid requirements of the current systems for a certain 'type' of voice ('specific characteristics required for voice'); the inability of the systems to change over time as 'changes in voice affect the system'; the memory and understanding requirements and difficulty for people with a 'condition that affects memory or cognition'; the requirement for patience (i.e. a 'lack of patience with system' being detrimental); and the requirement to 'learn the operation of the device'.

the frustrated person who raises his voice to command it, which is a natural instinct, actually ended up being worse and it was a discipline to remain monotonal and calm, which again is a contra-indication for somebody in an emergency situation

I mean he's been using that for a long time but he still forgets his curtain commands and he'll do an 'open' and it will set the intercom off and then he has to do 'rotate' I think. I can't remember what it is for his curtains or for his blinds and I would regard him as a good user but he still forgets some of the less frequently used commands. So there is a mental overload on certain aspects of it, as there is with any environmental control system.

Factors related to difficulty with the features of existing devices were also noted: the 'difficulty of interfacing the device with some peripherals'; the 'difficulty of training new functions' and 'limited menus' in particular.

Discussion

This paper has considered the perceptions of professionals involved in the provision of environmental control systems in the UK with regards to the provision of speech-driven environmental control systems. These data add to the existing anecdotal evidence around the reasons for the low provision rate of speech driven environmental controls. The reasoning around the provision of speech driven systems and the implications for professionals considering provision are discussed below. The data also offer some potential for helping to understand the empathy of professionals when considering users' needs and this is also discussed.

Clinical Relevance

Professionals 'triage' potential end-users for provision of speech-driven environmental control systems based on a 'mental model' of successful end-users, including considerations of the cognitive load and indicators based on the client's condition. Professionals also consider the possible service implications of providing a system and use the systems in fairly well defined scenarios as part of the range of environmental control solutions. Some examples of this include use of the systems as a 'last resort' and use for very able clients (e.g. those with a spinal cord injury) as an augmentation to their environmental control systems.

There appears to be a low rate of provision of speech-driven environmental control systems and some of the reasoning for this is identified in this paper; the main reason identified is the perceived unreliability of the systems. A

number of end-users of environmental control systems are heavily reliant on their systems, for example they may spend periods of time alone, use it to call for help or operate safety critical functions, thus the reliability of systems can be crucial. The perception of unreliability leads to a requirement for alternative back up systems where possible and a reluctance to use speech-driven systems except when absolutely necessary or for non-safety critical functions. When systems are provided they are provided with the implication of the possible unreliability carefully considered through risk assessment.

Professional Empathy

The results of this analysis can be compared against the published end-user data (Judge et al. 2009) in order to provide a basic evaluation of whether professionals and end-users perceive the same issues to be important. In Judge et al. (2009) twelve end-users of speech driven environmental controls were interviewed and the interview data were coded into a framework of themes and sub themes that described the range of issues considered by end-users. The main themes developed within this work were: factors influencing failure; factors influencing success; interface; speech driven environmental control usage; and background. Comparing the results shows that professionals discussed the majority of the sub-themes covered by end-users. This correlation of sub-themes is indicative of the professionals' understanding of and empathy with issues faced by end-users.

The variances in the compared sub-themes also have some interest. For example: professionals strongly considered an end-user's disability and cognition, however these were not considered strongly by end-users. End-users strongly considered 'feedback' from the system to be important but this

was not covered during the professionals' discussions. These two examples of variance may be explained through consideration of the different view-points of the end-users and professionals. In the case of cognition and disability, this is evidently a consideration for professionals when considering provision and forms part of their assessment procedure and 'mental model' of a successful user. End-users will be biased in the consideration of these issues and, since they are, by definition, successful users of the system, they will already have the 'pre-requisite' cognitive ability or a 'compatible' disability. In terms of the consideration of feedback, this may be because there is no choice of feedback with current systems and professionals may only need to consider the binary decision of whether the feedback would be suitable for an end-user or not.

'Factors influencing failure' was the theme that had the least similarity between the professionals' and end-users' perceptions. This may indicate that professionals have a 'positive mental model' when considering provision of speech-driven environmental controls – i.e. professionals consider provision according to criteria for success. Further work is indicated into comparing end-users' and professionals' perceptions to explore the clinical reasoning behind assistive technology provision.

Both end-users and professionals identified reliability as the most important reason for failure and this can be considered as the most important issue that needs to be overcome in the development of future speech-driven environmental control systems. The lack of reliability has already been shown to be detrimental to the use of speech-driven environmental control systems

and has also been shown in this paper to influence the decision making of assistive technology professionals. It is suggested that lack of reliability is the key reason for the low rate of provision of these systems.

Limitations of the study

The professionals' data could have been gained solely through interviews (as with the end-user data) which may have resulted in different and/or richer results. The decision to use a focus group to obtain this data was based on two initial interviews. On initially reviewing this data it was felt by the researchers that the interview format was not proving effective at fully examining the topic and stimulating in-depth consideration of the issues. The researchers felt that the focus group potentially offered a better method for exploring the topic with these professionals.

Participants in the focus group were likely to be regarded as 'expert' or 'enlightened' and thus may not be representative of the wider group of assistive technology professionals. This possible bias arises for two reasons: firstly, participants had been involved in providing a speech-driven system and secondly they were actively involved in the Assistech mailing list (through which they were recruited). The first source of bias was not possible to overcome due to the nature of the study, however it may have been possible to recruit more widely. There is, however, no centralised co-ordination of environmental control systems and thus wider recruitment is difficult at best, if not impossible.

Participants were also biased as being from a clinical-technical background i.e. no occupational therapists recruited. Again, this bias is likely to be

inherent due to the fact that these systems are considered as a 'last resort' and demand extensive support. The fact that the services replying to the recruitment email were all clinical-technical could possibly imply that non clinical-technical services are unlikely to consider such systems or are not as active on the Assistech mailing list. Further work could be carried out to evaluate the variances of provision based on the varying models of service delivery existing in the UK.

The data were analysed rigorously using a recognised qualitative analysis method, however the re-use of the existing framework, developed through analysis of end-user interviews on the same topic, did not provide potential for a fully open analysis of the data. The advantage of this approach, however, was the ability that this provided to allow comparison of the data from the two groups and to investigate the two different perspectives on the same topic.

Conclusion

This paper, based on analysis of data from a small focus group of assistive technology professionals, has provided evidence regarding the low provision of speech-driven environmental control systems in the UK. It has highlighted key reasons why professionals consider provision of such systems and suggests that they empathise with the challenges faced by end-users in the operation of these systems. The main issues highlighted were around reliability of the systems and this corresponds with the results of the end-user study on the same topic. These results suggest that, if reliability could be improved, speech driven systems could become more widely used and the

benefits, such as speed and compactness, could be appreciated by more endusers.

References

Craig, A. et al., 2005. The efficacy and benefits of environmental control systems for the severely disabled. Medical science monitor: international medical journal of experimental and clinical research, 11(1).

Geggie, C., 2003. Voice Control of Environmental Control Systems. ACNR, 3(4), 22-24.

Hawley, M.S. et al., 2007. A speech-controlled environmental control system for people with severe dysarthria. Med Eng Phys, 29(5), 586-593.

Holme, S. et al., 1997. The use of environmental control units by occupational therapists in spinal cord injury and disease services. Am J Occup Ther, 51(1), 42-48.

Judge, S. & Colven, D., 2006. Switch access to technology - A comprehensive Guide., The ACE Centre. Available at: http://ace-centre.hostinguk.com/assets/Product Downloads/SwitchScanningMaster_8_472.pdf.

Judge, S. et al., 2009. Speech-driven environmental control systems - a qualitative analysis of users' perceptions. Disability and Rehabilitation: Assistive Technology, 4(3), 151-157.

Maguire, S., McCann, J. & Swallow, M., 2001. An audit of the provision of environmental control systems in Northern Ireland, 1992-1997. Clin Rehabil, 15(3), 320-323.

McDonald, P. & St Leger, S., 1996. Provision of environmental control systems in the North West of England. J Public Health, 18(4), 443-448.

Novak, S.A., 1998. Environmental control systems - an audit of existing provision in three inner London districts. Clin Rehabil, 12(1), 88-93.

Palmer, P. & Seale, J., 2007. Exploring the attitudes to environmental control systems of people with physical disabilities: A grounded theory approach. Technology and Disability, 19(1), 17-27.

Paul, S. et al., 2006. Exploring environmental control unit use in the age group 10-20 years. Int J Ther Rehabil, 13, 511-6.

Ritchie, J. & Lewis, J., 2003. Qualitative Research Practice: A guide for Social Science Students and Researchers, Sage Publications Inc.

SHAW, H., 1962. 'Possum': an electronic aid to fuller living for the severely disabled. Physiotherapy, 48, 308-309.

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

Wellings, D.J. & Unsworth, J., 1997. Fortnightly review: Environmental control systems for people with a disability: an update. BMJ, 315(7105), 409-412.

Williams, M.H. & Bowie, C., 1993. Evidence of unmet need in the care of severely physically disabled adults. BMJ, 306(6870), 95-98.

Table 1: Participants

Professional	Job Title	Involvement in study
1	Service Manager	Exploratory Interview
2	Rehabilitation Engineer	Exploratory Interview
3	Clinical Technologist,	Focus Group
	Service Manager	
4	Clinical Technologist,	Focus Group
	Service Manager	
5	Clinical Scientist	Focus Group
6	Clinical Scientist	Focus Group
7	Clinical Technologist,	Focus Group
	Service Manager	
8	Service Manager	Focus Group