Older Adults and the Usability of Speech Interaction

Mary Zajicek and Andrew Lee
Department of Computing
School of Technology
Oxford Brookes University
Wheatley Campus
Oxford OX33 1HX
Tel: 00 44 1865 469935
Email: mzajicek@brookes.ac.uk

Richard Wales
Age Concern Oxfordshire
The Clockhouse, Long Ground
Oxford OX4 7FX
Tel: 00 44 1865 717614
Email: richardwales@ageconcernoxon.com

ABSTRACT
The main aim of this paper is to outline how a new technology, Voice XML can be used to provide Internet access for older adults who do not have a computer. The design of the Voice XML dialogues which enable the interaction is informed by the experience of workers on the Age Resource Desk at Age concern Oxfordshire and comments from their clients who experimented with a VoiceXML system. It is hoped that systems such as the ones described in this paper will to address the challenge of enabling older adults to participate in ICT. Older people often have little knowledge of computing and in addition age associated impairment particularly memory and sight loss make using standard desktop computers difficult. The new solution put forward here uses XML based technology to provide alternative forms of Web access through VoiceXML which offers Web access over the telephone and a grammar system from which to build dialogues. The design of the Voice XML dialogues is crucial to the success of the system. Although there is no need to learn how to use a computer, speech systems also pose some problems for older adults.

This paper describes special features in the dialogue which help older adults to use the system and reports experiments carried out to see how successful the dialogue was for older adults. It is acknowledged that the special design put forward in this paper does not help older adults alone. Features which help users to remember how to use a dialogue are useful for everybody.

Keywords
Older adults, Voice dialogues, Speech interactions, VoiceXML, Web accessibility, Learning ICT.

INTRODUCTION
This paper reports a synthesis of work carried out by tutors at Age Concern Oxfordshire and research work from the Speech Project at Oxford Brookes University. It describes the Voice Access Booking System (VABS) which enables clients at Age Concern Oxfordshire to organize their IT taster session by booking, canceling or rearranging them using a speech dialogue over the telephone from their own homes. Although telephone access to the Web for older adults promises widening accessibility and inclusion, its effectiveness for this user group depends on the usability of the VoiceXML dialogues themselves. In particular older adults find difficulty in remembering strategies at the interface therefore a successful dialogue will reduce the need to rely on memory.

The paper covers the particular characteristics of older adults which make them special. They do not form an homogeneous group, rather they display a wide diversity of ability and perspective, thus rendering standard User Centred Design methods unsatisfactory.

This paper gathers together the experience of tutors teaching older adults at Age Concern Oxfordshire, and then covers the design of dialogues for a Voice Access Booking System (VABS) and in particular explores ways in which instructions and context sensitive help can be incorporated into the dialogue design to help older adults.

The paper then describes the experiments carried out with the VABS dialogue to find out which parts of the dialogue were successful and which parts weren't. The users' paths through the dialogue were recorded and their divergence from the optimum path investigated, thus providing useful pointers to underlying principles in voice interaction design for older adults. The paper concludes by discussing how the design principles established with the VABS can be used in other contexts.

OLDER ADULTS' HCI ISSUES
Older adults form a significant proportion of the population numbering some 9.4 million, with predictions that there will be more over 65s than under 16s by 2014 in the UK [8], 'In 2036 it is predicted that there will be 39,000 people

over 100 years old' [1].Additionally almost half of the older population is aged at least seventy-five, with even larger relative increases being experienced in people over eighty-five.

Ageing can result in a combination of accessibility issues. Declines in sensory, perceptual, motor and cognitive abilities that occur with the normal ageing process have implications for interface design. Many of these are catalogued by Morrow and Lierer [7], whereas Morris, Czaja and Hawthorn [6], [3], [5] have described the different declines in abilities that occur with age and the implications of these for human-computer interface design. However very little research has been carried out into what makes an interface easy to use by older adults,

Recommendations From Age Concern Oxfordshire

Tutors at Age Concern Oxfordshire have identified five key factors which come into play when older adults learn how to use ICT, some of which contrast with the experience of younger people, whereas others are simply more extreme examples of the needs of any learners.

- many older adults cannot remember long instructions if given to them before they attempt a new operation. It is much better to give little hints as they are going along. Zajicek and Morrissey [10] performed experiments which measured the amount of information retained by learners from long and short messages, and found that older adults retained less information from long messages, and that listening to longer messages interfered with the ability to retain information. Similar experiments carried out with younger people showed that the length of message made no difference to information retention:
- all learners need time to assimilate new information, and some older learners, particularly those with poorer memory, less good vision and low dexterity take longer to adjust to the computer. However passing time is not an issue for many older adults therefore learning can take as long as necessary. The time and diligence which older people can invest in learning new skills are real assets;
- older adults, like all learners, will benefit from positive reinforcement as they learn how to interact with a computer. Zajicek and Hall [9] showed that with personal support in the form of a helper who simply answered yes or no to their questions older adults were able to use a Web browser where they had not been able before. At the Age Resource Desk the joy which clients have shown as a result of the support and positive reinforcement given to them has been a great reward for everyone;
- show and do, or learning by example is a very powerful form of teaching for older adults. This is important as some older adults only learn by watching others and so the level of interaction between them and the tutor is minimal. Showing how to do something is different from giving instructions. Show and do is difficult to arrange when people are using a voice system from home but points to the importance of demonstrating the system while potential users are in the company of trainers and other clients at IT taster sessions or an Age Resource Desk;
- no client will progress in any way unless they are comfortable with the equipment and set up. Time is spent at the Resource Desk setting up machines using the accessibility options to meet the needs of individual clients. This is a challenge for the introduction of VoiceXML technology which is designed to be used by people when alone in their houses.

Why Older Adults are Not a Homegeneous Group

As people grow older, their abilities change. This process includes a decline over time in the cognitive, physical and sensory functions, and each of these will decline at different rates relative to one another for each individual. This pattern of capabilities varies widely between individuals, and as people grow older, this variability increases. In addition, any given individual's capabilities vary in the short term due, for example, to temporary decrease in, or loss of, function due to a variety of causes including illness, blood sugar levels and state of awareness.

This collection of phenomena presents a fundamental problem for the designers of interactive systems, whether they be generic systems for use by all ages, or specific systems to compensate for loss of function. Current software design typically produces an artefact which is static and which has no, or very limited, means of adapting to the changing needs of users as their abilities change. Even the user-centred paradigm looks typically at concerns such as representative user groups, without regard for the fact that the user is not a static entity. Processes in use at present are ineffective in meeting the needs of many user groups, or addressing the dynamic nature of diversity. A method is required for designing for dynamic diversity. Gregor, Newell and Zajicek [4] put forward the case for a User Sensitive Inclusive Design approach which would actively seek out diversity and ensure that test users are recruited who represent the whole range of abilities, and which would promote an awareness when defining user requirements of a system that this diversity exists. This is particularly important when using questionnaires, as described in Section 5, where widely differing perceptions average out to a middle of the road score. In fact an average is probably meaningless when working with older people; the range of different perceptions is the reality.

THE SPEECH INTERFACE

In this section the focus shifts from reflections on teaching older adults ICT and the challenges of designing systems for this group, to the use of voice dialogues by older adults. The dialogue driven system under discussion, the Voice

Access Booking System (VABS) was built for Age Concern Oxfordshire, and was based upon a Web accessible database which holds the bookings for IT taster sessions at the Age Resource Desk.

Using the VABS the session organiser can manage the database of appointments using a standard XML based graphical interface and during office hours clients can book, cancel or re-schedule a session, which the organiser records on the database. The new feature being reported upon here is that clients can also phone up the database and make their own bookings using a VoiceXML dialogue which interacts with the database.

VoiceXML http://www.w3.org/TR/voicexml20/ is a new technology, which is only a couple of years old. It offers the dialogue builder simple building blocks known as form and menu, and a set of grammars. The challenge for the dialogue builder is to use these components to construct a successful dialogue, which older adults will be able to use unaided in their own homes to organise their own taster session appointments.

The VABS was built following discussion with Age Concern Oxfordshire and the establishment of the five key factors described in Section 2.1

The system was designed to allow the user to complete the following tasks:

- book a taster session with a reminder call;
- book a taster session without a reminder call;
- cancel a taster session;
- notify the database if they are going to be late.

Users are either guests or registered users who have provided their details to the system and who therefore can use the reminder call service.

USABILITY ISSUES IN SPEECH INTERACTION

The Design of Current Telephone Answering Systems

Current telephone answering systems with which we are all involved every day do not adhere to the principles that make for good design for older adults.. They use a hierarchical menu system, which places a high cognitive load on the user, as they are required to work out which option covers the goal that they have in mind. Messages are long and often contain preliminary instructions with many options to choose from.

Error recovery and help when a mistake has been made are frequently non-existent. If a user makes a mistake when choosing options it is difficult to get back to where they were. Clients at Age Concern Oxfordshire frequently mention that they do not like such systems.

Representation of the VABS Dialogue

The dialogues developed for the VABS are best-illustrated using flow diagrams. The key to the flow diagrams is shown in Figure 1. Unfortunately there is no space to present a flow diagram for the whole system. This paper therefore presents one example to demonstrate usability principles for older adults, and to that end Figure 2 represents the dialogue for setting a reminder call. In the following sections we pick out instances of the dialogue structures in Figure 2 which are shown to be useful for older adults.

Keep messages short

As described in Section 2 a major usability aim is to keep output messages or instructions as short as possible. The nature of the VABS dialogue makes for short messages, as users are questioned as they go through the dialogue rather than presented with lists of options as in a standard hierarchical menu driven system. The messages are courteous but hold at most one item of information or ask for only one item of data. One possible path through the dialogue in Figure 2 illustrates this as follows:

VABS – A call reminder can be set on the day prior to your booked session. What time do you want your call?

User – 11 o’clock

VABS – Please confirm that you want a call reminder at 11 o’clock

User – No

VABS – Try again to select a call reminder. What time do you want your call?

Embedding help in dialogues

The primary challenge for the dialogue designer is to provide context sensitive help and instructions to help the older adult to use the dialogue. As described in Section 2 another major usability aim is to keep output messages as short as possible, and provide positive feedback to users. The use of context sensitive help provides a way to support this aim.

Help Interrupt

When using the dialogue older adults were able at any time to say help and the system would jump to help instructions. Currently instruction takes the form of explanatory text relating to the current state as follows:

Area 1 – This system uses voice recognition to understand your commands that should be spoken clearly. Giving your name helps the system to locate your current sessions and gives you access to more functions. You need to contact the centre during office hours to register your name.

Area 2 – This system is designed to offer users the ability to book or cancel a computer taster session. Speak your commands clearly and try to use the words given in the question.

Area 3 – Sessions are generally available Monday to Friday. Please state only one of these days clearly.

Area 4 – Sessions run from 10:30 to 15:30 every hour. Please say something like 'ten thirty' to see if the session is available.

Area 5 – To cancel a session you need to find the day and time of the session and then confirm you wish to cancel it.

Area 6 – Guest bookings have to be located by day and time. Please ensure that you only cancel your guest booking.

Area 7 – Registered users can tell the centre they are running late. This helps the organisation of the taster sessions and your help is appreciated.

Area 8 – Registered users may ask to receive a telephoned reminder for a training session the day before the session starts. Every attempt will be made to call at the set time.

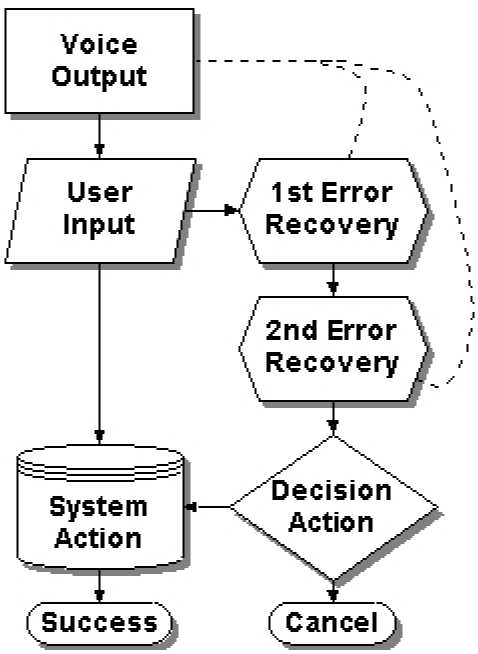

Figure 1 A generic error recovery dialogue

Embedded help/information

The issue in the call reminder task is how to let the user know that call reminders can only be set using the twenty-four hour clock and are possible only on the hour.

The information is treated as help information, which is embedded in the error recovery loop. A generic error recovery loop is shown in Figure 1. The main reason for this feature is to avoid lengthy introductory messages but also the probability of a user already knowing this information is high since they must be already registered in order to use the service. Guests cannot reach this point in the dialogue.

In effect the user is prompted through the dialogue. One client at Age Concern Oxfordshire commented that 'the dialogue takes what you have given and then prompts for the gaps'

Use of defaults

The dialogue fragment shown in Figure 2 features the use of the default message 'Unable to determine a time for a reminder call. Would 7 pm be OK?', which offers a possible retrieval of the task by offering a default booking time for a call rather than allow the user to leave the dialogue without having completed their task

This approach contrasts with that of a standard telephone answering system in which users often have to start the call all over again if they make an error or forget something.

Confirmatory Sentences

To make the user feel in control of their interaction confirmatory sentences were used which can be found at several points in Figure 2., for example. 'Please confirm that you want a call reminder at <time>', or 'Thank you. You will receive a call at your registered number at <time> the day before your session'. Defaults and confirmatory sentences provide positive reinforcement in line with the five key learning factors described above.

DETERMINING USER ATTITUDE

The Pre-test Questionnaire

A questionnaire based on a five point Likert scale was used at the start of experimentation to determine user's attitudes to the system. The effects of dynamic diversity described above, comes into play here, as averaging the numerical responses for many questions in the questionnaire would result in a rather bland mid number where the actual responses were decisive and very different. Subjects were encouraged to speak freely around each question and although average scores are not reported here verbal responses are recorded below:

Q1 – Access to information about sessions over the telephone will be useful.

All answers to this were positive.

Q2 – Access to my personal information, such as address, over the telephone will be useful.

Answers were divided for this question. Some users were concerned about who had access to this information.

Q7 – Making a booking for a training session over the telephone will be useful.

All answers were positive, which is encouraging as the whole development focuses on the booking of sessions.

Q8 – The ability to make contact out of hours will be useful.

A range of answers were received for this. Some users are busier than others so appreciate the increased ability to contact the centre, whereas others are happy to make contact during the day.

Q9 – A call to remind me about a booked taster session will be useful.

Mainly positive or neutral answers to this indicate that users are happy to organise themselves although may appreciate a timely reminder on some occasions, probably when they booked the session some time in advance.

Task	Optimum	Worst	User 1	User 2	User 3	User 4	User 5	User 6
1. Guest Main Menu.	9	18^{F}						9
2. Registered Main Menu.	5	14^{F}	7	5	7	7	5	
3. Guest booking, yes to call.	6 + 1	$16^{F F}$ + 1						10 + 1
4. Registered booking, yes to call.	6 + 4	$16^{F F}$ + 9^{F}	11 + 6^{F}	11 + 9	13 + 5		13 + 7	
5. Guest booking, no to call.	6	$16^{F F}$						
6. Registered booking, no to call.	6	$16^{F F}$				16		
7. Guest cancellation.	7	$12^{F F F}$						12
8. Registered cancellation.	1-4	1 - ∞		4	4			
9. Guest late.	1	1						1
10. Registered late.	4	6		5				

Table 1 Number of steps for each task, giving optimum, worst case and steps for each user by task.

Q10 – I'd prefer to speak to a human rather than a computer system.

Most users would rather speak to a human. This is not a reflection on his or her dislike for computer systems but everyone appreciates that a conversation with a human is more involving than a computer system, even though it may be less efficient.

Q11 – I would prefer the computer system to have a male/female voice.

No respondent expressed a preference for either voice. Clients felt that the voice they conversed with didn't matter as long as it was clear.

These responses indicate a positive and open-minded response to the idea of VoiceXML access. Given the negative attitude to standard telephone systems reported above these answers were encouraging.

Evaluating the Dialogue

Looking at users' paths

The dialogue of the VABS was tested with six users at Age Concern Oxfordshire. The set of tasks for the whole system were identified and the optimum and worst potential routes traced. A route includes system output, user input and error cycles. Fatal errors that return the user to the main menu or the operator are denoted by an '^{F}', one for each potential error.

Table 1. shows the ten possible tasks carried out in the VABS, together with the optimum and worst-case routes for each task. The route metric for each user is also shown. The figures in the table indicate that some tasks are better supported by the dialogue than others. Task 9: Guest late, for example is supported to such an extent that the user can carry it out in one step and cannot go wrong. Whereas with Task 8: Registered cancellation, the user could remain in a continuous loop. However Table 1 shows that the two users who tried Task 8 both carried it out in 4 steps. This demonstrates that the number of possible steps in a dialogue cannot be used alone as a usability measure. The quality and positioning of messages appears to be playing an important part in helping users to avoid the continuous loop situation.

The designer's aim is to reduce the number of potential steps and ultimately make them the same as the optimum path. This is particularly challenging for data entry tasks where input recognition quality is not easy to predict or control.

Five users were added to the 'names.g' file so that they could proceed as registered users, anticipated to be the most common status for users of the system. One user acted as a guest. The user tests were recorded on video camera and the script file from the simulation software stored for analysis. Table 1 details the tasks taken by each user and the number of nodes on the route visited to successfully complete the task. This is a useful way of documenting users pathways through the system as it enables easy inspection to see which dialogues are most successful in terms of the number of user steps.

Table 2 shows that no user filled in the booking form in the optimum number of steps, and only one fatal error was encountered throughout the tests which was due to a misunderstanding by the recognition engine of the desired time for a call reminder. Only one user asked for help and then proceeded to answer the next prompt successfully.

Task 3: Guest booking, yes to call, and Task 4: Registered booking, yes to call, both also shown in Figure 2., were the most problematic because they rely on voice recognition for data entry. The user paths for those tasks that did not involve data entry were much nearer to the optimum score. Design alternatives for the dialogues supporting Tasks 3 and 4 are put forward in Section 5.5.

What the users thought

All users expressed dislike for the voice used for speech synthesis which came with BeVOCAL Café, the VoiceXML platform that was used for dialogue implementation. It was American and was purposely set to speak slowly. Despite disliking the voice, users were perfectly able to understand the output.

User 1 – Expressed dislike for the slow and laborious American voice. Used full sentences when replying to the system or used leaders and trailers around the keywords such as 'please…', 'could I…'.

User 2 – Originally from America so happier with the accent, also seemed to enjoy better voice recognition. Picked out the use of keywords successfully although had a tendency to use leaders such as 'try…' and 'how about…'.

User 3 – An example run through performed solely by the wizard as the user didn't really want to talk to the system. The user could appreciate the use of the system out of office hours or on the days when the office is closed.

User 4 – Strong dislike for the voice. Became frustrated at the first loop and re-prompt at an error cycle. Not a user of the training sessions anyway and lives so close they would just come around!

User 5 – Used the help command on the main menu, although this may have been for exploratory purposes. Became disgruntled at the persistent use of their full name within the prompts, an issue that would be resolved by the use of the underlying database that uses forename and surname separately.

User 6 – Enjoyed the challenge of acting as a guest of the system. Listened closely to the voice and completed the booking form quickest of all. Understood that they would just need to pop into the centre to become a registered user.

All users appeared to get quickly frustrated when they encountered an error recovery cycle that failed to explain what they did wrong. The input they gave may have been valid but was not recognised or the input keywords 'Help' and 'Menu' did not convey enough information. This occurred most frequently when booking a session; all but one user gave the time as a particular hour (i.e. 11 o'clock) but the centre runs sessions at half past the hour. Many users that call the centre ask for a morning or afternoon session and the session organiser then allocates them a particular slot.

Reducing Input Errors

The data in Table 1. shows that most tasks were completed in near to the optimum number of steps and users were able to use the entire dialogue to complete their tasks. We can also see that the tasks involving user input by voice were the least successful. The research group intends to implement a dialogue design feature, also researched in the group, which was previously developed for answering systems for hospital and university reception. This feature would replace the user input with a 'yes', 'no' type of dialogue proposed by Brownsey, Zajicek and Hewitt [2] which will in effect perform binary chops on the possible entry data. For example when a time for a taster session is required, instead of entering the time, the user would be asked 'Would you like your IT session in the morning or afternoon?', as normally occurs in the human-human conversation when the session organiser sets up the session. If the answer is morning the system would then respond 'Before eleven o'clock or after?' The dialogue would continue to halve the search area until a time is found. The authors acknowledge that the binary chop will increase the number of questions required but on the other hand will reduce the number of input errors. Evaluation of this design in terms of optimum pathways will be blended with user satisfaction metrics to discover if this approach improves the dialogue over all.

CONCLUSIONS

The system described here demonstrates that it is possible for older adults to interact successfully with the Web using Voice XML. Particular attention was paid in the dialogue to support for those who find difficulty in remembering the sets of steps and strategies needed to operate interactive systems.

Older adults experience similar problems with products such as interactive TV, and automated machines, where the kind of spoken help and instructions embedded in the VABS dialogue could be very useful. As interaction with interactive products become more complex new research questions arise such as, when should help be activated? Should the system learn about a particular user and learn how to help them? Should help kick in when a non-optimal path through the interaction is detected? Should it be set to detect the point when users take a set of wrong paths in their interaction?

The ability to set up reminders is a powerful aspect of VoiceXML technology, and can be extended to pre-programmed reminders which place telephone calls to remind an older adult to take some medicine, switch on the heating or remember that a particular person will be visiting. Potentially then a remote carer can populate a database with reminders, which will prompt telephone calls at prearranged times, a useful development for people trying to organise their elderly parents for example.

Armed with an efficient set of electronic tools which replace the need for a good memory, older adults will be able to remain independent for longer.

ACKNOWLEDGEMENTS

The authors would like to thank the tutors and clients at Age Concern Oxfordshire for their valuable support for the work described above.

REFERENCES

1. Age Concern, Age Concern, General Statistics 2001,http://www.ageconcern.org.uk/AgeConcern/information_428.htm ,2001
2. Brownsey K. Zajicek M. Hewitt J., 'A structure for user oriented dialogues in computer aided telephony', *Interacting with Computers*, Dec., 1994, Vol. 6.4, 433 – 449, 1994.
3. Czaja S., Interface Design for Older Adults, In: A. F. Ozok and G. Salvendy (eds) *Advances in Applied Ergonomics*, USA Publishing, 262 – 266, 1996
4. Gregor, P, Newell, A, Zajicek, M, 'Designing for Dynamic Diversity – interfaces for older people'. Paper presented to the 5th ACM/SIGAPH Conf. on Assistive Technologies, Edinburgh, 2002.
5. Hawthorn D., Possible Implications of Ageing for Interface Designers, Interacting with Computers, Vol 12, 507- 528, 2000.
6. Morris J., User Interface Design for Older Adults, Interacting with Computers Vol 6, No 4, 373 – 393, 1994.
7. Morrow, D., Leirer, V., Ageing, Pilot Performance and Expertise, In: Fisk, A., & Rogers, W., (eds), Handbook of Human Factors and the Older Adult, Academic Press, pp 199 – 230, 1997
8. Office of National Statistics, Ageing Population, http://www.statistics.gov.uk/cci/nugget.asp?id=287, 2001.
9. Zajicek M. & Hall, S., 'Solutions for elderly visually impaired people using the Internet', HCI 2000, Sunderland, pp 299 – 307, 2000.
10. Zajicek M., Morrissey W., 'Speech output for older visually impaired adults', Proc. IHM-HCI 2001, pp 503 – 513, 2001.