USING SPEECH RECOGNITION FOR CHILD COMPUTER INTERACTION WHEN DEVELOPING SOFTWARE FOR YOUNG CHILDREN

Tony Nicol
University of Central Lancashire
Preston PR1 2HE
anicol@uclan.ac.uk

ABSTRACT

This paper provides an overview of a PhD project which investigates the feasibility of using commercial off-the-shelf speech recognition software as an input mode with young children. This research project also looks at the special requirements of children as computer users and proposes a set of Child-Computer Interaction design guidelines. The project also suggests methods of enabling an unreliable but developing recognition technology to be used effectively and demonstrates the effectiveness of these findings by developing and evaluating a speech-enabled educational computer application for young children.

Keywords

Child Computer Interaction, Speech Recognition.

1. INTRODUCTION

There are several areas in the education of young children which would benefit from the child being able to speak to the computer: reading practice is an obvious example. There are two key issues to be considered when developing a speech-aware application for young children: designing the child-computer interaction, and making effective use of speech recognition technology. Both areas have been studied during this PhD project, which has culminated in the development and evaluation of a simple reading aid. Interface design for infant or pre-reading children can be very challenging as the interface has to be very simple yet remain effective [1]. In general HCI design, much of the feedback to the user is textual and a considerable amount of the design effort concentrates on the location, size, colour and wording of the text [2]. However, the luxury of written text is largely unavailable to the designer of an interface for young children.

Most research in the field of HCI has concentrated on interface design for adults. This project aims to add to the growing body of knowledge in the area of interface design for children, with a particular emphasis on using speech recognition as an input mode. Several researchers have studied children speaking to a computer [3,4,5], but these research teams tend to use custom-built speech recognition software. There is little research into using commercial speech recognition technology with children. Some success with older children using commercial technology has been reported [6] but much research work has yet to be done with young children.

2. SPEECH RECOGNITION ACCURACY

Recognition accuracy only has to be reasonable for it to be usable for dictation, but in other instances, accuracy is paramount. Consider the situation where the word “help” is displayed for reading and child says “hello”. If the speech recognition system accepted this answer as correct, the application may respond with “well done”, which is clearly educationally unsound. It is the effective implementation of this type of application that is the focus of this PhD project.

Tests were carried out to determine the accuracy of several speech recognition systems. The outcome of these tests determined whether further research was reasonable and on which systems to concentrate [7].

Because it is very difficult to have regular access to children for long periods of time, an accuracy test environment was developed. This enables recordings

of speech from children to be collected to build a speech corpus that can be used to test the recognition accuracy of a custom application. It was found that the recognition accuracy for all the commercial systems under test was very poor when used with children. However, the accuracy improved significantly when the systems were trained. Training enables the speech recognition system to adapt to the speaker's voice. Unfortunately this requires the speaker to read a set text provided by the speech recognition system. But if the children could read the text, there would be little point in using a system to teach them to read. Tests also showed that after training, although recognition rates were reasonably high, the false recognition rates (where the speaker says one thing but the recogniser produces something different) were unacceptable for educational applications. This was the key problem to be solved if speech recognition was to be used effectively.

3. INTERFACE DEVELOPMENT

Several non-speech-enabled computer applications have been developed and tested with children to learn how to improve the child-computer interaction. Based on this knowledge and the discussions with teachers, a set of application design guidelines were proposed and a further application was developed to test the effectiveness of these guidelines [8]. Areas evaluated include the use of concatenated human speech to replace text for user feedback, of randomness to avoid pattern spotting, of interactive help [9] and of user profiles to enable children to progress at their own rate.

4. VOCABULARY TESTER

To pull together the results from the two research areas of speech recognition and child computer interaction, an application has been developed to enable children to practice their vocabulary. The National Curriculum strategy specifies that a child should be able to read on sight a specific set of words called "high frequency words" [10]. The application displays the words randomly and the child reads them. The application tries to recognize the speech and responds accordingly. A method has been developed to overcome the serious problem of false recognitions yet still provide a reasonably natural and intuitive user interface; after each utterance the application will ask the child to confirm what the recognition system thinks he said and therefore prevent a misrecognition from being accepted. This application is currently being evaluated in a primary school.

5. INTERIM CONCLUSIONS

The research has enabled the development of a more natural interface. Results have shown the application to work reliably and effectively when used in the classroom. However, the limitations of the technology need to be clearly understood. This thesis identifies the limitations and discusses methods to improve the effectiveness of speech input to help other researchers and software developers to produce effective speech enabled applications for young children.

6. REFERENCES

[1] Clements D. and Swaminathan S. (1995). Technology and school change: New lamps for old? Childhood Education, 71, 275 - 281

[2] Marcus A. (1992). Graphic Design for Electronic Documents and User Interfaces: Addison-Wesley.

[3] Oviatt S.L. (2000) Talking To Thimble Jellies: Children's Conversational Speech with Animated Characters, In B. Yuan, T. Huang & X. Tang (Eds.), Proceedings of the International Conference on Spoken Language Processing (ICSLP'2000), Vol. 3, (pp. 877-880). Beijing, China: Chinese Friendship Publishers.

[4] Mostow J., Aist, G., et al. (2003). Evaluation of an automated Reading Tutor that listens: Comparison to human tutoring and classroom instruction. Journal of Educational Computing Research, 29(1), 61-117

[5] Russell M.J., Series R.W., et al.(2000) "The STAR system: an interactive pronunciation tutor for young children", Computer Speech and Language, Vol. 14, Number 2, 161-175

[6] O'Hare E and McTear M. (1999). Speech recognition in the secondary school classroom: an exploratory study. Computers & Education, 33(1), 27 - 45.

[7] Nicol A., Casey C. and MacFarlane S. J. (2002) Children are Ready for Speech Technology - but is the Technology Ready for Them? Proceedings of Interaction Design and Children, Eindhoven.

[8] Nicol A and Casey C. CAL (2003) interface design for young children: a case study in literacy: CAL'03 Proceedings.

[9] Nicol A. Help! What do I do now? Interactive Help for Young Children - A Case Study. (2002) Proceedings of Comp@uclan. Preston, UK.

[10] Department for Education and Employment. (1998) The National Literacy Strategy Page 61. ISBN 085522 714 1.