

Accepted for presentation at the International Society for Augmentative and Alternative Communication (ISAAC) Conference, Barcelona, July 2010; <http://www.isaac2010.org/>

Further Development of the PhonicStick: The application of phonic-based acceleration methods to the speaking joystick

Authors: Ha Trinh, Annalu Waller, Rolf Black, Ehud Reiter

Abstract

The PhonicStick is a novel Augmentative and Alternative Communication joystick-like device which enables individuals with severe speech and physical disorders to access forty-two sounds (i.e. phonics) and blend the sounds together to create spoken words. The device aims to allow the users to select phonics and produce speech without the need for a visual interface. One of the problems with the current prototype of the PhonicStick is that the phonic entry is relatively slow and may involve many physical movements which will cause great difficulties for users with poor hand function. Therefore, in this research we are investigating whether natural language processing (NLP) technology can be utilized to facilitate the phonic retrieval and word creation processes. Our goal is to develop a set of phonic-based NLP acceleration methods, such as phonic disambiguation and phonic prediction, which will reduce the user effort required to select the target phonics and improve the speed of producing words. This paper will discuss the challenges of applying such methods to the PhonicStick and report on the current state of the development of the proposed techniques. The presentation will also include a live demonstration of the latest prototype of the PhonicStick.

Keywords: PhonicStick – Literacy Learning – Acceleration Techniques

Extended Abstract

BACKGROUND

Over recent years there has been an increasing number of speech generating devices (SGDs) developed to provide a “voice” for individuals with complex communication needs (CCN). However, most of these devices are symbol or letter-based and require the users to navigate a visual interface to generate spoken output. To address the problem of symbol or letter-based retrieval, our research group has been developing a novel speaking joystick, known as the PhonicStick, which provides the users with access to the 42 spoken phonics introduced in the Jolly Phonics literacy program (Lloyd, 1998) and allows them to combine the phonics into words without the need for visual cues (Black et al., 2008). The device is designed to be used both as a communication aid and a literacy aid to support teachers and speech and language therapists in teaching literacy to children with severe speech impairments. As a communication aid, the device will allow individuals who have not mastered literacy skills to create novel words and messages. As a literacy aid, it will enable learners to play with sounds and become familiar with the sound structure of the language. This will assist individuals with CCN who are incapable of sounding words out and at risk of not being literate in developing their phonemic-awareness, an essential prerequisite for reading (Blischak, 1994).

To evaluate the potential of the joystick interface for phonic-based communication, we undertook a pilot study in which a prototype speaking joystick was developed to enable access to a subset of six phonics, including /s/, /t/, /a/, /p/, /i/, and /n/. An evaluation carried

out with seven children who have varying degrees of physical and communication disabilities demonstrated that the participants were able to recall the phonic positions and generate novel words without using a visual interface (Black et al., 2008).

Following the positive feedback on the PhonicStick gathered from the pilot study we are aiming to extend the prototype to include into the device all 42 phonics. These phonics are divided into eight groups which can be accessed by moving the joystick from the centre of the joystick workspace towards eight compass rose directions.

One of the major problems with our current prototype speaking joystick is that the phonic entry and word creation are usually very slow. To produce a given word the users will need to select all phonics that make up the word. The selection of each phonic usually consists of three stages: (1) Push the joystick from the centre of the joystick workspace into the group to which the target phonic belongs; (2) Move the joystick around its circumference to find the target phonic; (3) Move the joystick back to the centre to select the target phonic. Therefore, the process of creating a given word may require a considerable number of physical movements. Furthermore, the feedback collected from the pilot study showed that users with poor hand function had great difficulties with the second stage of the phonic selection (i.e. moving the joystick along circular paths), which further slowed down the process. Thus, the purpose of our research is to investigate whether acceleration techniques can be incorporated into the device to enhance the phonic entry and reduce the physical workload needed for the users to generate words.

PHONIC-BASED ACCELARATION TECHNIQUES

Background

Acceleration techniques are commonly used in AAC systems to improve the user input performance by reducing the number of keystrokes required to produce a given word or message. These techniques are particularly beneficial for physically disabled users whose each movement is not only slow but also difficult and tiring. Amongst various acceleration methods that have been incorporated into the AAC devices, we are focussing on the disambiguation and prediction strategies as they are frequently used in the orthographic-based AAC systems.

Disambiguation techniques are often applied in devices with an ambiguous keyboard (i.e. the keyboard on which each key represents a group of characters or symbols), such as 12-key mobile phones. Disambiguation can be performed at the character level to predict which of the possible characters on the pressed key is actually the user's intended character. If the prediction is incorrect, the user can repeatedly press the "error" key until the correct character is selected. Previous research has shown that character disambiguation can result in approximately 90% efficiency (i.e. 90 characters produced by 100 key presses) for a 12-key keyboard (Arnott, 1992). Disambiguation can also be applied at the word level to estimate the N-character word that the user has intended to type given a sequence of N key presses. A well-known implementation of the word-level disambiguation is the T9 feature on mobile phones (James, 2000).

Prediction techniques are another method used to achieve keystroke savings in AAC devices. A prediction system presents the user with a list of either letters or words that are most likely to be entered following the previous input. The user will then scan the prediction list to select the desired letter or word. It has been reported that the use of word

prediction can lead to 48% keystroke savings (Higginbotham, 1992). Furthermore, prediction programs are also beneficial for users with poor spelling skills (Glennen, 1997).

Challenges

There are two major challenges in applying such acceleration methods as disambiguation and prediction to the PhonicStick. The first challenge is that most currently available disambiguation and prediction programs are letter-based rather than phonemic-based. Although word prediction features have been introduced in several phoneme-based assistive systems, such as the REACH Sound-It-Out Phonetic Keyboard™ (Schroeder, 2005), none of these systems work with phonic input. There is also no existing system that provides phonic-based disambiguation functions. There is little or no previous research on whether it would be feasible to adapt the letter-based or phoneme-based acceleration techniques to allow for the phonic input. The second challenge is that prediction programs usually display a list of letter or word suggestions on a visual interface. The PhonicStick is aimed to work without such visual cues. Thus, our problem is to investigate other feedback channels (i.e. other methods of presenting the list of predicted phonics and words to the user) to be used in the disambiguation and prediction processes.

Proposed Solutions

We are developing a set of phonic-based disambiguation and prediction features, as outlined below:

- **Phonic disambiguation**
Using this feature the program will attempt to guess which of the possible phonics is required each time the user moves the joystick into a phonic group and place it at an easy-to-access location. The phonics in that phonic group may be reordered based on the probability that the user will select those phonics. Different methods of arranging phonics into groups will be investigated to improve the efficiency and usability of the disambiguation function.
- **Word disambiguation**
This feature will allow the user to create a complete word without having to access the exact location of the phonics that make up the word. The user will only need to specify to which phonic group a phonic belongs. The program will then estimate which word should be produced given the sequence of selected phonic groups. If there are several candidate words matching that sequence, the program can speak those words out so that the user can select their desired word from the spoken word list by pressing a 'select word' switch.
- **Phonic prediction**
This function will produce a list of probable next phonics after each phonic entry. We are currently investigating two methods of presenting the predicted list to the users without the need for a visual display. The first method makes use of the auditory feedback channel, i.e. the predicted phonics will be spoken out. The second method uses the force feedback channel in which physical forces will be applied to push the joystick towards the location of highly probable next phonics. The idea of combining prediction with force feedback is adopted from an existing dynamics and probabilistic text entry system (Williamson, 2003).

Our aim is to incorporate these features into the PhonicStick and evaluate them with our user group to identify the pros and cons of each feature. Both quantitative and qualitative

data will be collected from the user feedback to assess the efficiency and appropriateness of the proposed features.

CONCLUSION

In this paper we have described the problems in the development of the acceleration techniques for the phonic-based speaking joystick and proposed the solutions for those problems. During the presentation we will provide the latest update on the implementation of those proposed solutions.

ACKNOWLEDGEMENTS

This research is supported by the Scottish Informatics and Computer Science Alliance (SICSA) and the School of Computing, the University of Dundee.

REFERENCES

- Arnott, J. L., Javed, M. Y. (1992). Probabilistic character disambiguation for reduced keyboard using small text samples. *Augmentative and Alternative Communication*, 8, 215-223.
- Black, R., Waller, A., Pullin, G. & Abel, E. (2008). Introducing the PhonicStick: Preliminary evaluation with seven children. *In: 13th Biennial Conference of the International Society for Augmentative and Alternative Communication (CD ROM)*, Montreal, Canada.
- Blischak, D. M. (1994). Phonologic awareness: Implications for individuals with little or no functional speech. *Augmentative and Alternative Communication*, 10, 245-254.
- Glennen, S. L., Decoste, D. C. (1997). *The Handbook of Augmentative and Alternative Communication*, Thomson Delmar Learning.
- Higginbotham, D. J. (1992). Evaluation of keystroke saving across five assistive communication technologies. *Augmentative and Alternative Communication*, 8, 258-272.
- James, C., Longe, M. (2000). Bringing text input beyond the desktop. *In: Conference on Human Factors in Computing Systems*, The Netherlands.
- Lloyd, S. M. (1998). *The Phonics Handbook*, Chigwell, Jolly Learning Ltd.
- Schroeder, J. E. (2005). Improved Spelling for Persons with Learning Disabilities. *In: The 20th Annual International Conference on Technology and Persons with Disabilities*, California, USA.
- Williamson, J., Murray-Smith, R. (2003). Dynamics and probabilistic text entry. Dept. Computing Science, University of Glasgow.