

A Hindi Virtual Keyboard Interface with Multimodal Feedback: A Case Study with a Dyslexic Child

Yogesh Kumar Meena
Intelligent Systems Research
Centre, Ulster University
Derry~Londonderry, UK
meena-y@ulster.ac.uk

Anirban Chowdhury
Centre of Mechatronics
Indian Institute of Technology
Kanpur, India
anir@iitk.ac.in

Ujjwal Sharma
Department of Humanities and Social
Sciences, Indian Institute of Technology
Kanpur, India
usharma@iitk.ac.in

Hubert Cecotti
Department of Computer Science
College of Science and Mathematics
Fresno State University, USA
hcecotti@mail.fresnostate.edu

Braj Bhushan
Department of Humanities and Social
Sciences, Indian Institute of Technology
Kanpur, India
brajb@iitk.ac.in

Ashish Dutta
Centre of Mechatronics
Indian Institute of Technology
Kanpur, India
adutta@iitk.ac.in

Girijesh Prasad
Intelligent Systems Research
Centre, Ulster University
Derry~Londonderry, UK
g.prasad@ulster.ac.uk

Up to 15% of the Indian school-going children suffer from dyslexia. This paper aims to determine the extent to which existing knowledge about the eye-tracking based human-computer interface can be used to assist these children in their reading and writing activities. A virtual keyboard system with multimodal feedback is proposed and designed for a lexically and structurally complex language and optimized for multimodal feedback involving several portable, non-invasive, and low-cost input devices: a touch screen, an eye-tracker, and a soft-switch. The performance was evaluated in terms of text-entry rate, information transfer rate, and type of errors with three different experimental conditions: 1) touch-screen condition with auditory feedback 2) eye-tracking condition with auditory and visual feedback, and 3) eye-tracking and soft-switch condition with auditory and visual feedback. The proposed multimodal feedback has shown a significant improvement in text-entry rate with less error. This work represents the first virtual keyboard with multimodal feedback for dyslexic children in the Hindi language, which can be extended to other languages.

Virtual keyboard; graphical user interface; eye-tracking; multimodal feedback; dyslexia.

1. INTRODUCTION

Dyslexia is characterized by significant difficulty with speed and accuracy of decoding word, spelling and text comprehension. Studies have found the effect of text presentation on reading performance in dyslexic cases (Rello et al., 2012). Letter reversal is a common characteristic of dyslexia. Eye movement during reading has been used as a robust indicator of deficiencies seen in dyslexics. It can be used to record and understand moment-to-moment cognitive process during reading (Rayner, 1998). Reading necessitates visual scanning to

decode the written word which, in turn, is affected by various linguistic factors such as the word length and frequency. Researchers have used behavioral measurements (e.g., reaction times, response accuracy and visual function) and electrophysiological measures (e.g., event-related potential, visual evoked potential) to understand underlying factors in developmental dyslexia that can be exploited for training purposes.

Researchers have examined auditory and visual processing in children with dyslexia (Wright & Conlon, 2009) to understand the problems genesis associated with it. According to Stein (2001)

dysfunction of the magnocellular pathway (M-pathway) is responsible for the visual, auditory and motor problems related to dyslexia. Magnocells respond to the onset and offset of stimuli (Macknik & Livingstone, 1998) and the anomalies are found in magnocells in the Lateral Geniculate Nucleus and the Medial Geniculate Nucleus. Given the understanding of anatomy related to dyslexia, one can assume that training modules involving visual and auditory feedback should prove to be more effective.

Showman and Ahissar (2006) argue that visual deficit is not the cause of reading difficulties experienced by dyslexics; rather it is the consequence. A large number of studies have combined auditory and articulatory training and found improvement in phonological deficit in such children (see Joly-Pottuz et al., 2008). However, they found motor tapping task as 'the best predictors of training efficiency' (see Joly-Pottuz et al., 2008, pp. 402). The lack of integration of the auditory-visual system has been reported in dyslexic children (Breznitz, 2002). Although Breznitz (2002) studied visual and auditory modalities, he examined them separately. Thus, it is difficult to draw an inference as to how the cross-modal processing operates. Moreover, these studies do not shed light upon relevant motor problems related to dyslexic children. These research findings motivated us to incorporate visual and auditory feedback in a single interface and evaluate its efficacy in dyslexic children.

Studies endorse oculomotor deficits in children with developmental dyslexia (Bucci et al., 2008). Studies using visual evoked potentials (Kubova, et al., 1996), behavioral measures of visual function (Slaghuis & Ryan, 2006) and eye-tracking (Bellocchi et al., 2013) have an important role in helping understand, develop, and test training protocols for children with dyslexia. For languages read from left to right the first fixation point is usually between the beginning and middle of the word (Li, et al., 2011). Studies suggest that dyslexics are characterized by higher number of, and longer fixations during reading a word, pseudoword, or sentence. In addition, they show shorter saccades with increased regressions. They also fixate only once on few words or even skip them; while on more number of words they show multiple fixations (Hawelka et al., 2010). Furthermore, the words length affects the gaze duration (Luca et al., 2002).

It is important to note that the findings pertaining to oculomotor deficits in dyslexic children are based on Western studies on English, Spanish, German and other languages, but not Hindi. Although Hindi is also read from left to right, it has peculiar characteristics such as the usage of diacritics (*matras*) and killer strokes (*halants*). In view of the above, we developed and evaluated a virtual keyboard interface with visual and auditory feedback

for dyslexic children. Having developed a multimodal device with real-time feedback features, we tested its efficacy on a dyslexic participant. The present case study explicates the usage of the virtual keyboard interface with multiple feedbacks, and their significance in errors committed while typing Hindi text.

2. SYSTEM OVERVIEW

The developed text entry system consists of three main components (Figure 1): (i) an application interface with a graphical user interface (GUI) representing a Hindi virtual keyboard, (ii) a multimodal character entry system, wherein three different input devices: a touch screen, an eye-tracker, and a single input switch, were implemented with an application interface to type a text, and (iii) a multimodal feedback system, wherein two different multimodal feedbacks (auditory and visual) are designed and integrated with the application interface for users. The developed virtual keyboard application comprises of two main parts; the first part represents the ten possible commands and the second part is an output text display where the user can see the typed text. This virtual keyboard application interface is adopted from Meena et al., 2018. The detailed description of choosing an alphabetical organization with script specific arrangement layout, tree structure of the ten commands, and working of application interface is explained in Meena et al. (2018). In particular, the tree-based structure of the GUI provides the ability to type 45 Hindi language letters, 17 different *matras* (diacritics) and *halants* (killer strokes), 14 punctuation marks and special characters, and 10 digits (0 to 9). Other functionalities to edit the text, such as delete, delete all, new line, space and go back commands for corrections are included.

On a virtual keyboard using eye-tracking, it is necessary that the user is given an efficient feedback that the intended command box/character was selected in order to avoid mistakes during copy spelling and increase efficiency. Hence, two types of visual feedbacks are added to the application—positive (for intended item selection) and negative (for accidental item selection). The feedback was represented by a change in the colour of the button border, from silver to green (positive feedback) or red (negative feedback) depending upon maintenance of gaze. When the user fixates and maintains his/her gaze to a particular button for a time duration t , the colour of the button border changes linearly in relation to the dwell time Δt and the border becomes greener (more red) in the positive (negative) feedback. The visual feedback allows the user a continuous adjustment and adaptation of his/her gaze to the intended region on the screen. An audio feedback is provided to the user through an acoustic beep after successful

execution of each command. This sound makes them proactive so that they can prepare for the next command. To improve the system performance by using minimal eye movements, the last five typed characters are displayed in the GUI under each command box, helping the user to see the previously written characters without shifting significantly their gaze from the desired command box to the output display box (Meena et al., 2018; Cecotti, 2016).

Three different modalities were integrated with the virtual keyboard application (Figure 1). First, the touch screen facility for the user wherein a user can type any Hindi language sentence by simply touching the required character on the laptop screen. With the touch screen facility only auditory feedback is provided to the user as an acoustic beep after a successful execution of the command. Second, an eye-tracker facility was integrated for the user wherein user can type any Hindi language sentence by pointing the gaze on the required character on the screen. In this modality, the command selection was achieved with a dwell time set to 1.5 s ($\Delta t=1.5$ s) (Meena et al., 2018; 2017; 2016). With eye-tracking facility only, the auditory and visual feedback are provided to the user. In the third modality, the eye-tracker and a soft-switch were integrated wherein the eye-tracker was used for pointing to the item, which was selected by pressing the soft-switch. This condition provides auditory and visual feedback to the user.

3. METHODS

3.1 Participant

One male 15-year-old dyslexic volunteer participated in this study. The participant had no vision correction and did not have prior experience with the application. He was selected from a special school in India. He fulfilled the criteria for Specific Learning Disability (SLD) screening. The Wechsler Intelligence Scale for Children (WISC: Wechsler, 1949) and NIMHANS Index of Specific Learning Disability (Kapur et al., 1991) were administered for measuring the Intelligence Quotient (IQ). He had significant difficulties in visual memory along with other academic dimensions of reading, writing, spelling, comprehension, and arithmetic. The dominant writing hand of the participant was right. The study was conducted following the Helsinki Declaration of 2000.

3.2 Procedure and Experimental Paradigm

The participant was initially assessed using the two psychological tools. Thereafter, he was comfortably seated in front of a laptop screen (DELL, 15.6 inches, 60 Hz refresh rate, optimum resolution 1920x1080, touch-screen). The distance between

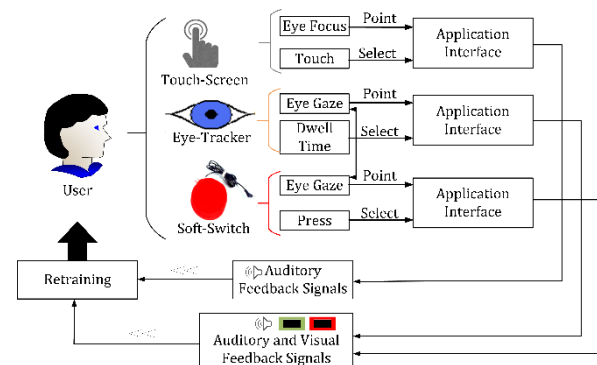


Figure 1: Block diagram of a proposed system with multimodal feedbacks

the participant and the laptop screen was about 80 cm. A portable eye-tracker (The Eye Tribe Aps, Denmark) was used for pursuing the eye gaze of the participant. Prior to each experiment, a calibration session lasting about 20 s, using a 9-point calibration scheme was conducted. The soft-switch was used as a single-input device to select a command on the screen (Singh & Prasad, 2015).

The typing task used in the experimental protocol involved 10 predefined Hindi words in increasing order of difficulty, based on the number of letters (two to six: घर, नमक, बचपन, जलमहल, and इधरउधर) and same number of letter with one extra matra (diacritic: राम, इमली, तरबूज, किसतरह, and संगमरमर). The participant was asked to type these words (i.e., copy spelling). The typed word was displayed in the output text box on the screen if it matched with one of the predefined words. Errors, if any, were saved in the log file without being displayed on the screen. The participant was asked to look at the word while typing. The time windows given to the participant for the first five words were 40 s, 60 s, 80 s, 100 s and 120 s, respectively and for the last five words were 60s, 80s, 100s, 120s and 140s, respectively. The time window was chosen as 20 s per character (letter or *matra*). Each predefined word appeared at the bottom of the screen one-by-one for these time durations. If the participant completed the task within the time frame then a message “बहुत बढ़िया” (well done) flashed in the message box, otherwise “अच्छा प्रयास” (well tried”) appeared in the box.

The participant performed the task by using the three different input modalities: a touch screen, an eye-tracker alone, and an eye-tracker coupled with a soft-switch, which provided three different experimental conditions (see Figure 1). The touch-screen condition, which represents a common input method for computing devices that is familiar to the participant was used as a baseline to measure the change in performance from switching from a touch screen to another modality that can include visual feedback. For each condition, only the correctly spelled characters were displayed in the output text display window, whereas, incorrectly spelled

characters were not displayed but saved in the log file for later performance evaluation.

4. RESULTS

The virtual keyboard with multimodal feedback was evaluated for all the three conditions across the ten difficulty levels (DLs). For computing statistical significance, the Wilcoxon signed-rank test was applied using false discovery rate (FDR) correction method for multiple comparisons on performance indexes across the DLs for each condition. The typing performances for all the three conditions (touch-screen (TS) condition with only auditory feedback, the eye-tracking (ET) condition with auditory and visual feedback, and the eye-tracking and soft-switch (ETSS) condition with auditory and visual feedbacks) are computed.

The average text entry rate with ETSS condition (12.83 ± 2.12 letters/min) is significantly superior ($p=0.017$) to the ET (9.91 ± 1.96 letters/min) and TS (9.37 ± 3.07 letters/min) conditions. Moreover, the average text entry rate with ET condition is also significantly higher ($p=0.037$) than the TS condition. When the performance was measured in terms of Information transfer rate (ITR) for command level (ITR_{com}) and letter level (ITR_{letter}) for each condition, the average ITR_{com} with ET (60.69 ± 5.24 bits/min) condition was significantly higher ($p=0.002$) than the ETSS (55.63 ± 20.09 bits/min) and TS (50.14 ± 8.07 bits/min) conditions. ITR_{letter} with ET (18.68 ± 13.29 bits/min) was also significantly higher ($p=0.027$) than the ETSS (13.88 ± 3.95 bits/min) and TS (14.83 ± 8.60 bits/min) conditions. In addition, the average time (ms) to select the commands was higher with the TS (3362.88 ± 2990.09 ms) condition than ET (2975.27 ± 951.73 ms) and ETSS (2364.67 ± 1019.99 ms) conditions. It suggests that participant took more time to select the commands with TS condition (auditory feedback) than ET and ETSS conditions (visual and auditory feedback).

Message (whether task is completed or not) and errors were also recorded in log file across the DLs for each condition. The participant received a message 'well-done' across the DLs for each condition when he completed the task within the pre-defined time. Furthermore, the errors were measured for each DLs and for each condition. The participant made a total of eight errors while completing the TS condition task. In particular, the participant made one error while completing the task with difficulty level 3 (task type बचपन). He first attempted व instead of ब character. Two errors were recorded with difficulty level 5 (task type इधरउधर). First, character घ was attempted instead of character ध and second, the go-back command was used while completing the task once. Two errors

were recorded with difficulty level 6 (task type राम). In this case, matra ॉ was attempted instead of matra ा and go-back command was once used by the participant. Further, with the difficulty level 7 (task type इमली), one error was recorded wherein go-back command was used in order to complete the task. The last two errors were associated with the difficulty level 8 (task type तरबूज), wherein character ज was attempted instead of character ञ and go-back command was used once. A total of three errors were made by the participant in order to complete the task with ET condition. One error was associated with difficulty level 6 (task type राम) and similar to the TS condition matra ॉ was attempted instead of matra ा. Two errors were recorded with difficulty level 8 (task type तरबूज) where ठठ characters were attempted twice instead of ज. Finally, we analysed the errors associated with the ETSS condition and found that the participant committed only one error. While attempting difficulty level 7 (task type इमली) participant attempted matra ि instead of ी.

5. DISCUSSION AND CONCLUSION

This case study substantiates the virtual keyboard application as an effective interface for cross-modal processing. The findings show some patterns as far as reading and writing of Hindi words are concerned and the analyses of text entry rate, ITR for letter level (ITR_{letter}) and command level (ITR_{com}) provide useful information for addressing some of the fundamental issues in the study of cross-modal processing in dyslexic children for Hindi language.

Most importantly, this study shows the efficacy of the virtual keyboard application. There was a systematic decrease in the number and nature of errors depending on the type of feedback: higher with only auditory feedback using TS, gradual reduction with auditory-visual feedback using ET only, and further reduction with the auditory-visual feedback using ETSS. Breznitz (2002) has referred to the lack of auditory-visual system integration in dyslexic children, while studying visual and auditory modalities separately. However, our study examined visual and auditory modalities both separately as well as integrated together. This study thus proved suitability of the multimodal virtual keyboard in Hindi for a child with dyslexia. This is encouraging; the system however needs to be tested on a large cohort of children with dyslexia to test the efficacy of the system as well as the nature of the best feedback. Currently, we are focusing on testing this system on a large cohort of dyslexic children.

REFERENCES

- Bellocchi, S., Muneaux, M., Bastien-Toniazzo, M. & Ducrot, S. (2013). I can read it in your eyes: What eye movements tell us about visuo-attentional processes in developmental dyslexia. *Research in Developmental Disabilities*, 34, 452-460.
- Breznitz, Z. (2002). Asynchrony of visual-orthographic and auditory-phonological word recognition processes: An underlying factor in dyslexia. *Reading and Writing*, 15, 15-42.
- Bucci, M. P., Bremond-Gignac, D., & Kapoula, Z. (2008). Latency of saccades and vergence eye movements in dyslexic children. *Experimental Brain Research*, 188, 1-12.
- Cecotti, H. (2016). A multimodal gaze-controlled virtual keyboard. *IEEE Trans. Human-Machine Systems*, 46 (4), 601-606.
- De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye movement study of developmental dyslexia. *Brain and Language*, 80, 617-626.
- Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition*, 115, 367-379.
- Joly-Pottuz, B., Mercier, M., Leynaud, A. & Habib, M. (2008). Combined auditory and articulatory training improves phonological deficit in children with dyslexia. *Neuropsychological Rehabilitation*, 18, 402-429.
- Kapur M, John A, Rozario J, Oommen A, Uma H. (2002). Assessment of specific learning disabilities. In Hirisave U, Oommen A, Kapur, M. (eds), *Psychological Assessment of Children in the clinical Setting*. Bangalore: Vinayaka Offset Printers.
- Kubova, Z., Kuba, M., Peregrin, J., & Novakova, V. (1996). Visually evoked potential evidence for magnocellular system deficit in dyslexia. *Physiological Research*, 45, 87-89.
- Li, X., Liu, P., & Rayner, K. (2011). Eye movements guidance in Chinese reading: Is there a preferred viewing location? *Vision Research*, 51, 1146-1156.
- Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, 1, 144-149.
- Meena, Y. K., Cecotti, H., Wong-Lin, K., Dutta, A. & Prasad, G. (2018). Toward Optimization of Gaze-Controlled Human-Computer Interaction: Application to Hindi Virtual Keyboard for Stroke Patients. *IEEE Trans. Neural Syst Rehabil Eng*, 26 (4), 911-922.
- Meena, Y. K., Cecotti, H., Wong-Lin, K., & Prasad, G. (2017). A multimodal interface to resolve the Midas-Touch problem in gaze controlled wheelchair. *IEEE Conf. on Engineering in Medicine and Biology Society (EMBS)*, 905-908.
- Meena, Y. K., Cecotti, H., Wong-Lin, K., & Prasad, G. (2016). A novel multimodal gaze-controlled Hindi virtual keyboard for disabled users. *IEEE Conf. on Systems, Man, and Cybernetics*, 3688-3693.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rello, L., Kanvinde, G., & Baeza-Yates, R. (2012). Layout guidelines for web text and a web service to improve accessibility for dyslexics. In *Proc. W4A '12*, Lyon, France, 2012. ACM Press.
- Shovman, M. M., & Ahissar, M. (2006). Isolating the impact of visual perception on dyslexics' reading ability. *Vision Research*, 46, 3514-3525.
- Singh, J. V. & Prasad, G. (2015). Enhancing an eye-tracker based human computer interface with multi-modal accessibility applied for text entry. *Int. J. of Computer Applications*, 130 (16), 16-22.
- Slaghuis, W. L., & Ryan, J. F. (2006). Directional motion contrast sensitivity in developmental dyslexia. *Vision Research*, 46, 3291-3303.
- Stein, J. (2001). The sensory basis of reading problems. *Developmental Neuropsychology*, 20(2), 509-534.
- Wechsler, D. (1949). Wechsler intelligence scale for children. New York: The Psychological Corporation.
- Wright, C. M. & Conlon, E. G. (2009). Auditory and Visual Processing in Children With Dyslexia. *Developmental Neuropsychology*, 34, 330-355.