Enhancing an Eye-Tracker based Human-Computer Interface with Multi-modal Accessibility Applied for Text Entry

Jai Vardhan Singh
Intelligent Systems Research Centre (ISRC),
University of Ulster,
Magee Campus,
N. Ireland, UK

Girijesh Prasad
Intelligent Systems Research Centre (ISRC),
University of Ulster,
Magee Campus,
N. Ireland, UK

ABSTRACT

In natural course, human beings usually make use of multi-sensory modalities for effective communication or efficiently executing day-to-day tasks. For instance, during verbal conversations we make use of voice, eyes, and various body gestures. Also effective human-computer interaction involves hands, eyes, and voice, if available. Therefore by combining multi-sensory modalities, we can make the whole process more natural and ensure enhanced performance even for the disabled users. Towards this end, we have developed a multi-modal human-computer interface (HCI) by combining an eye-tracker with a soft-switch which may be considered as typically representing another modality. This multi-modal HCI is applied for text entry using a virtual keyboard appropriately designed in-house, facilitating enhanced performance. Our experimental results demonstrate that using multi-modalities for text entry through the virtual keyboard is more efficient and less strenuous than single modality system and also solves the Midas-touch problem, which is inherent in an eye-tracker based HCI system where only dwell time is used for selecting a character.

Keywords
Multi-modal, eye-tracker, HCI, eye typing, eye-gaze.

1. INTRODUCTION
When we use gaze or eye for text typing (or entering), it is commonly referred to as gaze typing or eye typing. Its importance was felt mainly for people suffering from motor neuron disease, muscular dystrophy, strokes or spinal cord injuries [1], for whom eye typing maybe providing the only means of communication with their surroundings. Initially it was felt that using eyes for pointing purposes is not very fruitful [2] for normal users except for disabled users who don't have other options. However, this type of system can also be useful in situations where user's both hands are occupied and he/she needs another mode to assist or complete (execute) the task. In a comparative study, Murata [6] found out that middle age and older people felt more comfortable with eye-gaze input than mouse. An Eye-gaze system also helps older people who have problems due to the reduced motor activity, arthritis or tremors, they take less time in pointing with eye-tracker than mouse. A survey in [7] provides a detailed insight into areas where eye-tracker is used or can be used along with other related research works.

Gaze can be used as an input mode (or method) but it is not very convenient mode for control purposes, as eyes have naturally evolved for perceiving (or perception). Gaze is used to acquire information about the environment. Recently, eyes are also reported to be used for controlling, by dwell time,

blink, wink, and eye gesture [12]. Gaze broadly consists of saccades and fixations [3]. During fixations (200-600 ms) which are followed by saccades (referred to as a ballistic movement), we usually can view the objects (of interest). Normally during interaction with computers a user gazes different objects on the screen (first) before any action. If we use (add) this gazing mode with our input system, the whole input process may become very fast.

Eye-gaze is faster than any other pointing device for pointing purposes [5]. So if eye-gaze along with hand is used for interacting or inputting into a computer, the whole process will become more natural and faster. It also takes less time to train for becoming good (efficient) user of this system. However, the problem with gaze devices is that they are not as accurate as other devices like mouse, due to the size of fovea which restricts measured point of gaze.

There are situations referred to as locked-in-syndrome [13], when users can't move their eyes in all directions, for that scanning [16] method can be used so any system developed should support the scanning.

Although there is more than thirty years of research history in gaze typing, but still it has some inherent problems like false selection of letters i.e. Midas touch [3], fixing of head during usage [29] and losing eye tracker system accuracy during usage (run), which require system to be recalibrated. Dwell time is widely used in eye tracker based systems for selecting or executing objects on screen. Commonly used dwell time values are between 400 ms to 1000 ms [33, 21, 8]. Smaller values of dwell time leads to false selection, while with larger values, it becomes quite tiring and unnatural for eyes as eyes need to blink frequently for lubrication, which provide them relaxation. There are some recommendations [21] for best dwell time. Researchers also recommended adjustable dwell time value [21, 22] instead of constant dwell value, to achieve faster typing speed.

2. RELATED WORK
Researchers are continuously trying and proposing different approaches to overcome the Midas touch problem, which is a major cause of errors during gaze typing [18] and reducing its usability among normal users. A brief review of some popular approaches follows. In Manual and Gaze Input Cascaded (MAGIC) technique [2], researchers proposed two approaches for cursor control by eyes: (i) Liberal approach; (ii) Conservative approach. In *liberal approach* cursor appears near to the object where the user is presently looking at, if he/she wants to select the target, he/she has to move the cursor by hand to the object. While in *conservative approach* cursor does not always wait near to the gazed object, it reaches near

to the gazed object when user actuates the manual input device. For the object selection user has to move the cursor manually over to the object.

Researchers in [10] proposed a text writing system named 'Dasher' and they claimed that a user can write up to 25 words/min with it, after an hour of practice. User can reach the desired letter by navigation. When he/she reaches the desired letter (or box), the system shows next possible letters (or combination of letters). Authors also claimed that their system has less error rate than an on-screen keyboard; this is due to the zooming facility. This system is fast due to its ability to predict next letter(s), a user may write.

In the '*EyePoint*' system [11], researchers used some keyboard keys (for single click, double click, right click, mouse over, or start click & drag) instead of dwell time for selection, calling them hotkeys. They called this procedure, a two-step progressive refinement process, wherein a user first looks at a desired object or screen area, then presses a hotkey, again looks at the screen area which is now magnified and if he/she wants to select or execute an action then he/she has to release the hot key. Here they thus used a magnifying technique to overcome eye tracker accuracy problem.

Recently researchers in [37] used a tooth-clicker device for the selection of an object. They used tooth-clicker along with an eye tracker for text entry through an on-screen keyboard and compared their results with dwell time based eye tracker typing. With tooth-clicker, rate of incorrectly typed characters was higher than that with dwell time. It takes more time to learn the usage of tooth-clicker. Researchers suggested that tooth-clicker is not suitable for typing tasks as it requires frequent selection (click), rather it is suitable for tasks which are not selection intensive like web browsing or reading.

Authors proposed an approach in [43] named context switching to overcome the midas touch problem. Here they are replicating the virtual keyboard in two distinct regions of screen, calling it contexts. To focus the desired letter short eye fixations was used and switching context was used for selection. They reported that after 8 sessions users obtained the typing speed of about 12 words per minute (wpm).

Actually there is a need of system which may be used with patients having higher level of motor disability like amyotrophic lateral sclerosis [ALS] at its different stages [1] or related patients without need of much modification for each stage, as patients are usually busy in coping up of their illness physically & mentally. The system should take very less time to learn, easy to install or setup and must be used with very little effort. The existing systems require sufficient time for learning to use them and require more physical effort or attention to complete the task.

3. RELATED WORK – MULTI-MODAL HCI SYSTEMS FOR TEXT ENTRY

Use of multi-modality inputs may make user experience more natural. Nature has provided humans the capability to speak, listen, smell, look, touch, taste and feel and it is know that in day-to-day task completion, human beings use their multiple senses; for example during conversation people usually use voice, gestures and eyes at the same time. Multi-modality makes the system more accurate and robust [24] which may not be possible with single modality system. In [23] authors discussed in detail, practical, biological and mathematical reasons which are compelling us to opt for multi-modality over single modality. Adding multi-modality to a HCI system

makes it more tolerant to errors and easier to rescue from errors [23], [24].

In Speech Dasher [44] researchers used speech and direction of gaze for text entry with modified version of Dasher. In this system user has to first speak what he wishes to write then he has to correct the errors showed by the speech recognizer by gazing. They obtained typing speed of 40 wpm with expert users after four hours of practice.

In one study [45] researchers proposed a multi-modal system where they used eye gaze and speech for text entry. They incorporated multi-modal functionality in popular word processor. In their system user can issue a command by speech after gazing at desired key on onscreen keyboard for selecting it. They obtained mean typing speed 0.2 to 0.3 characters per second.

In one multimodal system SpeeG [46] where researchers used speech and body gestures to input text. This system targeted the text input for game consoles, set-top boxes and media centres. This multimodal system is using the modified Dasher [10] interface. They achieved the average text input speed of 6.52 wpm.

Authors in [47] introduced a new voice typing system where utterances of users are transcribed as they produced. By doing this errors in written text can be identified in real-time. This system was tested with touch screen devices. User can use the marking menu by applying touch gestures to correct the errors. This system has lower cognitive load and lesser corrections compare to dictation as reported by authors.

Multimodal systems developed mainly require speech, gesture etc. for fulfilment (execute) of the task, as it is known that speech and other bodily movements are severely affected in patients of motor neuron diseases [1]. So there is a need of multimodal system which uses very less body effort and can be operated by any body part which may remain active (at their illness stage) for fulfilment of task.

In this paper, using an eye-tracker system by Arrington Research, development of a novel multi-modal input device involving an eye-tracker and a soft-switch is reported. It facilitates a jitter -free cursor control. The eye-tracker system is integrated with the soft-switch, typically representing a 2-state input device. Through trials on a set of healthy individuals a comparative evaluation is made between the uni-modal eye-tracker operation and the operation of a multi-modal eye-tracker soft-switch combination, when used for text-entry in an on-screen virtual keyboard developed in-house. The experimental results clearly demonstrate the effectiveness of the multi-modal system from the point of view of typing rate and helps in resolving Midas-touch problem along with more comfortable to use.

4. METHODOLOGY
4.1. Participants
In our experiment four volunteers had participated, all were male between 28-40 years, all had normal vision. Three participants had no previous experience with eye tracker. All volunteers were regular user of computer systems.

4.2. Development of a Multi-modal System
In our study we used an eye-tracker device *ViewPoint EyeTracker ®* by Arrington Research [38]; it is head fixed version. An LCD monitor of 21" was used. We used an in-house developed virtual keyboard [28], shown in figure 1. This virtual keyboard (VK) has been selected because it has a

multi-level scanning facility and it is known that severely motor disabled persons may not be able to move their eyes in all directions, for them scanning is a preferred method, although typing speed through scanning maybe slow. VK used by us has optimised placement of its keys [28] based on frequency of occurrence of English alphabets in text, which take less time in typing during *scanning* mode in comparison with alphabetic layout keyboard. This VK has multiple modes access facility like *eye tracker, soft switch* or *brain computer interface (BCI)*; particular mode(s) can be used depending upon the patient's level of disability. As a representative of another modality, a *soft switch* along with USB switch interface used here looks like as shown in figure 2 and figure 3. This is a binary switch, typical characteristics of assistive devices for people with neuro-muscular disabilities, e.g., two-state brain-computer interfaces. This *soft switch* enclosed in a foam, is covered by a removable and washable velvet case. This switch can be easily connected to any USB port of a computer system, which *doesn't require any additional power & software to install*. The *soft switch* can be operated by finger, hand, foot, or head with very little effort, as it is developed as an assistive tool for disabled people. Any other switch with which user feels comfortable with our system can be used, from a range of switches [39] available. We implemented a cursor control algorithm in such a way that our cursor movement should remain smooth like a mouse movement (without jittering), for this we investigated several clustering and smoothening algorithms along with filtration techniques. Our eye controlled cursor is similar to mouse and we can control any application by eyes easily. You can use any virtual keyboard with our eye gaze system if typing speed is your main concern and you are not using it with severely motor disabled people.
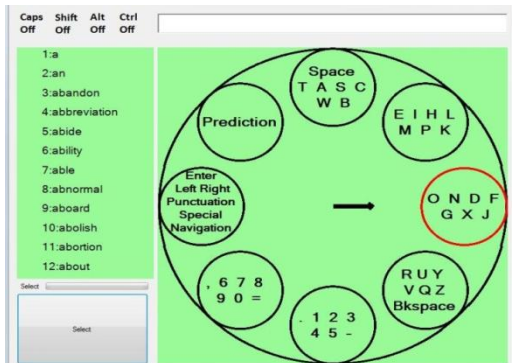
the character with the eye-tracking alone (i.e. typical eye typing) while during the *eye & soft switch* mode finding the block containing the character should be realized by the eye-tracking and the selection should be carried out by pressing the soft switch.

4.4. Experimentation Design
As mentioned above, experiment was divided into two parts:
(i)   Eye only
(ii)   Eye & Soft switch
and each part has three recording sessions for three different values of dwell time. An automated log was generated for each recording. Typing speed and errors are reported here. During typing we had not used word prediction or letter prediction [29] facility and virtual keyboard used by us is different from that by Mackenzie et al. [29], in which the number of keystrokes required per character (KSPC) [31, 32] affects the typing speed. The virtual keyboard used here has a minimum KSPC value of 2. Ideally it should be 1 [31].

5. RESULTS & DISCUSSION
Subjects were asked to type the sentence without any practice session. They also had to complete the both parts of the experiment with 10 minutes of break between them.

The grand mean of time taken for typing, grand average of typing speed and grand mean of *total error rate* [32] along with standard deviation are displayed in table 1 for the *eye only* mode and table 2 for *eye & soft switch* mode. For error calculation we used formula provided by Soukoreff et. al in [32] called *total error rate*.

Figure 1: Virtual Keyboard used in our study

Figure 2: Soft switch [40]

Figure 3: USB Switch Interface to connect soft switch to PC [41]

4.3. Procedure
A brief introduction about the experiment was given to volunteers before the start of the experiment. Just before starting a typing session, a calibration was carried out for each participant. We had not given practice session to participant to ascertain how comfortable and accurate they are with our system. A single 42 characters long sentence, which has all 26 English alphabets occurred at least once, was given to them for typing in two modes: *eye only* and *eye & soft switch*, for different dwell times. The sentence they have to type was : "*The quick brown fox jumps over a lazy dog.*". The sentence was shown only on top of the screen and we had not given time for practicing & memorizing it. No mistake in the finally typed sentence was allowed; if any mistake occurs they had to correct it. Participant can type either in lower case or upper case with proper punctuation characters. During the *eye only* mode, participants had to find and select (by using dwell time)

$$Total\ Error\ Rate = \frac{INF+IF}{C+INF+IF}\ X\ 100\%\ \ \ \ \ [32]$$

C- Correct characters in the transcribed text

IF- Number of errors made but corrected

INF- Number of errors made but not corrected

We have also shown results by excluding the data of the subject-3, who reported during subjective feedback taken after experiments that he was tired, due to which he was finding difficult to concentrate. We reported his results to show that being tired he committed less mistakes during the *eye & soft switch* mode for 3000 ms dwell time and more mistakes at 1000 ms dwell time, which are mainly double entry errors. Actually with 3000 ms dwell time, the character selection is almost dwell time independent, due to large time period, i.e. it has very little effect on selection of character blocks, as

usually subjects select (click) the block quite earlier than 3000 ms (with soft switch). This shows that multi-modal (eye & soft switch) mode helps in reducing error (or more error tolerant) at *extreme conditions* which often maybe the case with many users specially disabled ones due to their physical and mental state.

5.1. Analysis of Mean Time Taken

During the *eye tracker only* control mode (or uni-modal mode) *minimum mean time* taken for typing was 5.99 min when the dwell time was 800 ms (see table 1). *Maximum mean time* taken was seen when the dwell time was set to 1500 ms.

When subjects used *eye tracker and soft switch* control mode for typing, *minimum mean time* taken was 4.23 min at 3000 ms dwell time (see table 2). Here we observed that all subjects except one (about whom we already mentioned above), have taken mean time for typing, 3.16 min. They also committed fewer mistakes during typing. Here *minimum string distance* (MSD) error rate [32] is 0 in both modes.

subjects committed more mistakes, this was mainly due to *double clicks* i.e. one click by soft switch and another simultaneous click by dwell time reaching. Thus by setting the dwell time to a high value (as seen in dwell time value of 3000 ms, which is actually requirement for making system multi-modal otherwise it will be dwell time dependent uni-modal system), its effect on the character block selection becomes negligible hence false selections will not occur. This experiment also showed us that for multi-modal mode we have to consider dwell time value greater than 3000 ms to avoid errors caused by *double click* (or simply off the dwell time as it is not required due to soft switch).

5.2. Analysis of Typing Speed

Maximum grand mean typing speed of 7.01 characters/min obtained during *eye tracker only* control mode is at 800 ms dwell time. It is also evident from figure 4 that as the value of dwell time is increased, typing speed decreases.

When subjects used *eye tracker and soft switch* control mode, *maximum* grand mean typing speed obtained was 9.92 characters/min at 3000 ms dwell time. At 3000 ms dwell time, the group of subjects except the subject-3 obtained grand

Table 1. Mean values of time taken in typing during *Eye only mode* for different dwell time along with avg. typing speed and mean of total error rate.

S.No	Dwell Time (ms)	Mean of Time Taken (min)	Standard Deviation(ơ)	Avg. Typing Speed (Characters/ min)	Standard Deviation(ơ)	Mean of Total Error Rate (%)
1	1500	8.60 / 8.48[a]	1.17 / 1.40[a]	4.88 / 5.05[a]	0.77 / 0.91[a]	1.14 / 1.52[a]
2	1000	6.45/ 6.63[a]	1.18 / 1.38[a]	6.51 / 6.53[a]	1.22 / 1.44[a]	2.25 / 3[a]
3	**800**	**5.99** / 5.89[a]	1.62 / 1.97[a]	**7.01 / 7.67[a]**	2.06 / 2.45[a]	**1.72 / 1.52[a]**

[a] Excluding Subject-3

We have considered different sets of dwell time values in two modes. In *eye tracker only* mode we have selected 800 ms, 1000 ms & 1500 ms. The dwell time value of around 800 ms is recommended for novices. Dwell time values less than this are for experienced users so we had selected other two dwell time values 1000 ms & 1500 ms. During *eye tracker & soft switch* mode we don't need dominance of dwell time so we had considered comparatively high values of dwell time here i.e. 3000 ms, 1500 ms & 1000 ms.

Frequent false selections were observed widely during *eye tracker only* control mode, (figure 4), which was a major reason for longer time taken during typing.

mean typing speed of 13.36 characters/min which was much higher than 7.67 characters/min of the *eye only* mode's

best performance. Figure 6 shows best performance of both modes along with errors. We are getting quite interesting information from figure 5 that, when we decrease the value of dwell time, user starts committing more mistakes and his typing speed goes down which is just the opposite case of eye tracker only mode (cf., figure 4).

A phrase typing usually takes more time than simple character typing and we have given subjects a long sentence to type which also checked subject's endurance during using this system. Subjects' typing speed obtained at first attempt (also

Table 2. Mean values of time taken during *Eye & Soft switch mode* in typing for different dwell time along with avg. typing speed and mean of total error rate.

S.No	Dwell Time (ms)	Mean of Time Taken (min)	Standard Deviation(ơ)	Avg. Typing Speed (Characters/ min)	Standard Deviation(ơ)	Mean of Total Error Rate (%)
1	**3000**	**4.23 / 3.16[a]**	2.16 / 0.28[a]	**9.92 / 13.36[a]**	3.99 / 1.22[a]	**2.25 / 0.78[a]**
2	1500	5.80 / 4.95[a]	2.37 / 2.03[a]	7.25 / 9.59[a]	4.09 / 4.17[a]	8.48 / 7.76[a]
3	1000	10.29 / 6.19[a]	8.92 / 4.34[a]	4.08 / 8.85[a]	5.04 / 4.44[a]	22.91 / 12.94[a]

[a] Excluding Subject-3

During multi modal (eye tracker & soft switch) control, less false selections (mistakes) were observed (cf., figure 5) also, when the dwell time was increased to 3000 ms, two subjects didn't commit any false selections but at 1500 ms & 1000 ms

without practice session) shown in figure 7 demonstrates clear improvement in performance during *eye-tracker & soft switch* mode over the *eye tracker only* mode, which is quite encouraging.

The results of a study conducted on 12 subjects byusing Dasher system in [9], showed grand mean text entry speed of 2.49 wpm in first session, which increased significantly to 17.26 wpm in tenth session (after 2.5 hours of practice), this is a clear effect of learning. Another study in [19] also had found an increase in typing speed from first session to later sessions with other typing applications along with decrease in error rates.

We are showing results of the first session and most importantly we haven't used word prediction facility. It is a well-known fact that word prediction [29] increases typing speed by inserting desired word with fewer key selections. Researchers also worked on letter prediction [29] and found that the letter prediction is as good as word prediction and in some instances it is better. We are sure if we

subjects committed fewer or no mistakes at 3000 ms dwell time (cf., figure 5). If we increase the value of dwell time significantly (i.e. more than 3000 ms) or off the dwell time in our *multi-modal system*, by doing this we can remove false selections or midas-touch problem because it is related to dwell time. If midas-touch problem is solved, overall performance of our eye tracker system will be improved as we know that maximum time user usually spends on the interface for information purposes only (and due to which false selections occur).

5.4. 5.4. Less strenuous than single modality system

If we use a single modality system (eye tracker control only system), first the user has to find the desired character by gazing on screen and after that he/she has to gaze (dwell) at it
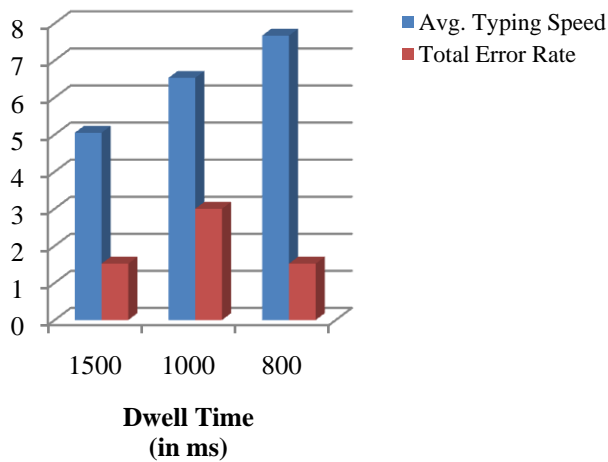

Figure 4: Result for Eye tracker Only mode; displaying average typing speed and total error rate at different dwell times.
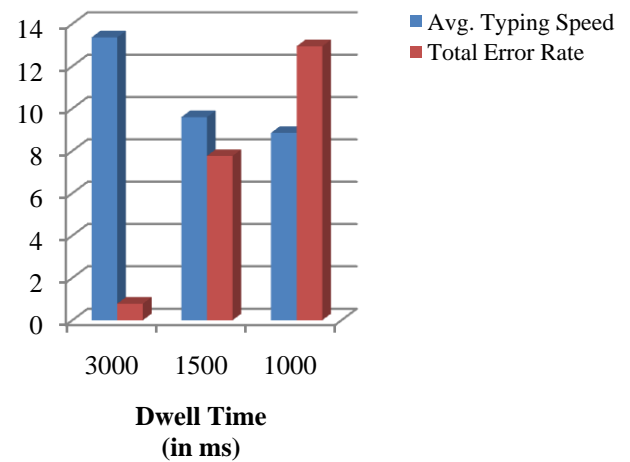

Figure 5: Result for Eye tracker with Soft Switch mode; displaying average typing speed and total error rate at different dwell times.

allow users to be more familiar with the system and enable word prediction facility with undo option of virtual keyboard, typing speed will increase significantly and errors will be reduced further.

5.3. Solving Midas-Touch problem

We have seen that by making our system multi-modal,

for pre-defined dwell time for selecting it. Gazing at single point for certain duration of time is strenuous task for eyes because naturally our eyes are used to blink regularly for comfort and lubrication. Applications like *typing*, which requires regular gazing for longer duration for the completion of task are very strenuous [11] for the eyes and neck when
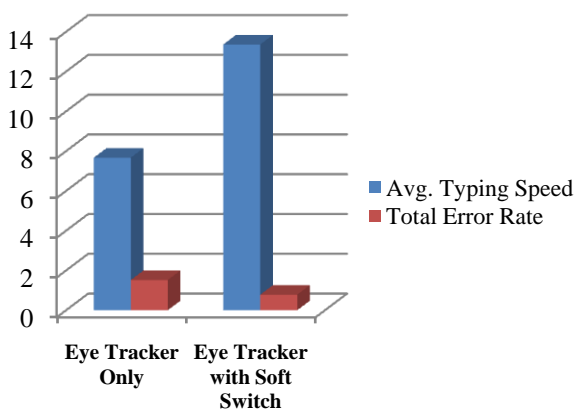

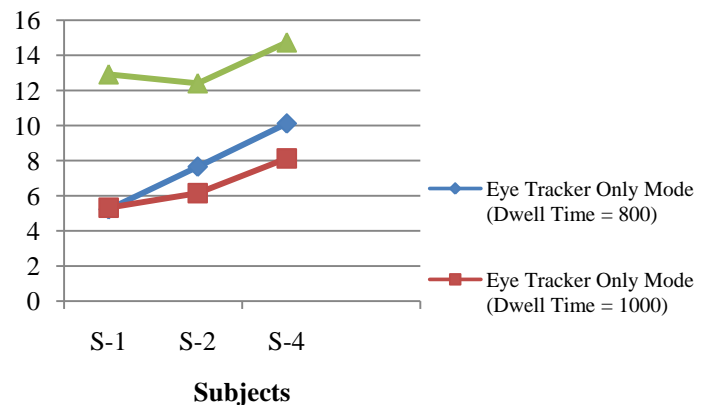**Figure 6: Comparing Best Results in both modes with *total error rate*.**

Figure 7: Typing speed of subjects category–wise

used with *eye-tracker only* system, so multi-modal system like us will provide great relief and ease. Even in Dasher also user has to gaze continually during zooming and selection of letter [9, 10].

In our *multi-modal system* we don't have to gaze for selection (click) of objects rather we are using *soft switch* for selection, which makes user free from unnecessary strain on eyes and unintentional selection. Here user can type simply by gazing at screen by his/her convenience without worrying about false selections or spend time in understanding or simply looking at interface, which provides users more freedom. Users need not be *very cautious* all the time during its usage as usually in case of *eye tracker only* system demands which will also relieve him/her from eye & mental strain.

Our system is not restricted to text entry only; you can use any other application with it also (considering the fact that eyes can't control very small icons or buttons by using present technology, so that application should not have very small icons).

6. CONCLUSION

The system proposed by us is very easy to setup and use. Multi-modal mode has shown significantly improved performance from the point of view of the typing speed and mistakes committed. This performance was achieved without any training session; usually training is desirable in this type of systems, which clearly indicates usefulness of this system. When user becomes familiar with our system overall performance will improve further. Text entry speed will increase significantly here, if you use word prediction and a virtual keyboard (VK) having an all key layout, where KSPC should be 1. Here I want to add that by using VK which has a value of KSPC equal to 1 alone, text entry speed will be double of the speed we achieved in our experiments because VK used by us has minimum KSPC equal to 2. But scanning facility of this VK is very useful for disabled users who have restricted eye movements due to their illness.

There is a need of more user trials with healthy and physically challenged users to prove its statistical validity.

7. REFERENCES

[1] Johansen A S and Hansen J P 2006 Augmentative and alternative communication: the future of text on the move *Universal Access in the Information Society* **5** 125-49

[2] Zhai S, Morimoto C and Ihde S 1999 Manual and gaze input cascaded (MAGIC) pointing *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems CHI '99 (Pittsburgh)* (New York: ACM) pp 246-53

[3] Jacob R J K 1991 The use of eye movements in human-computer interaction techniques: what you look at is what you get *ACM T. Inform. Syst.* **9** 152-69

[4] Hansen D W and Ji Q 2010 In the eye of the beholder: A survey of models for eyes and gaze *IEEE Trans. Pattern Anal. Mach. Intell.* **32** 478-500

[5] Sibert L E and Jacob R J K 2000 Evaluation of eye gaze interaction *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems CHI '00 (Hague)* (New York: ACM) pp 281-8

[6] Murata A 2006 Eye-gaze input versus mouse: Cursor control as a function of age *International Journal of Human-Computer Interaction* **21** 1-14

[7] Duchowski A T 2002 A breadth-first survey of eye-tracking applications *Behavior Research Methods, Instruments, & Computers* **34** 455-70

[8] Majaranta P and Räihä K -J 2002 Twenty years of eye typing: systems and design issues *Proc. of the 2002 Symp. on Eye Tracking Research & Applications ETRA'02 (New Orleans)* (New York: ACM) pp 15-22

[9] Tuisku O, Majaranta P, Isokoski P and Räihä K –J 2008 Now Dasher! Dash away!: longitudinal study of fast text entry by Eye Gaze *Proc. of the 2008 Symp. on Eye Tracking Research & Applications ETRA '08 (Savannah)* (New York: ACM) pp 19-26

[10] Ward D J and Mackay D J C 2002 Fast hands-free writing by gaze direction *Nature* **418** 838 (*arXiv preprint cs/0204030*)

[11] Kumar M, Paepcke A and Winograd T 2007 EyePoint: practical pointing and selection using gaze and keyboard *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems CHI '07 (San Jose)* (New York: ACM) pp 421-30

[12] Istance H, Bates R, Hyrskykari A and Vickers S 2008 Snap clutch, a moded approach to solving the Midas touch problem *Proc. of the 2008 Symp. on Eye Tracking Research & Applications ETRA '08 (Savannah)* (New York: ACM) pp 221-8

[13] Chapman, J E 1991 Use of an eye-operated computer system in locked-in syndrome In *Proc. of the Sixth Annual International Conference on Technology and Persons with Disabilities CSUN'91(Los Angeles)*

[14] Majaranta P, MacKenzie I S, Aula A and Räihä K -J 2006 Effects of feedback and dwell time on eye typing speed and accuracy *Universal Access in the Information Society* **5** 199-208

[15] Bee N and André E 2008 Writing with your eye: A dwell time free writing system adapted to the nature of human eye gaze *Perception in Multimodal Dialogue Systems* vol 5078, ed E André et al (Berlin Heidelberg: Springer) pp 111-22

[16] Shein, G F 1997 Towards task transparency in alternative computer access: selection of text through switch-based scannjng *PhD diss., University of Toronto*

[17] Majaranta P and Räihä K –J 2007 Text entry by gaze: Utilizing eye-tracking *Text Entry Systems: Mobility, Accessibility, Universality* (San Francisco: Morgan Kaufmann Publishers Inc) pp 175-87

[18] Hansen J P, Tørning K, Johansen A S, Itoh K and Aoki H 2004 Gaze typing compared with input by head and hand *Proc. of the 2004 Symp. on Eye Tracking Research & Applications ETRA '04* (New York: ACM) pp 131-8

[19] Pannasch S, Helmert J R, Malischke S, Storch A and Velichkovsky B M 2008 Eye typing in application: A comparison of two systems with ALS patients *Journal of Eye Movement Research* **2** 1-8

[20] Bolt R A 1980 Put-that-there: Voice and gesture at the graphics interface *Proc. of the 7th Annual Conf. on Computer Graphics and Interactive Techniques SIGGRAPH '80* (New York: ACM)

[21] Špakov O and Miniotas D 2004 On-line adjustment of dwell time for target selection by gaze *Proc. of the Third*

Nordic Conf. on Human-Computer Interaction NordiCHI '04(Tampere) (New York: ACM) pp 203-6

[22] Majaranta P, Ahola U -K and Špakov O 2009 Fast gaze typing with an adjustable dwell time *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems CHI '09 (Boston)* (New York: ACM) pp 357-60

[23] Sharma R, Pavlovic V I and Huang T S 1998 Toward multimodal human-computer interface *Proc. IEEE* **86** 853-69

[24] Oviatt S 2003 Advances in robust multimodal interface design *IEEE Comput. Graph. Appl.* **23** 62-8

[25] Kaur M, Tremaine M, Huang N, Wilder J, Gacovski Z, Flippo F and Mantravadi C S 2003 Where is it? Event synchronization in gaze-speech input systems *Proc. of the 5th Int. Conf. on Multimodal Interfaces ICMI '03 (Vancouver)* (New York: ACM) pp 151-8

[26] Oviatt S, DeAngeli A and Kuhn K 1997 Integration and synchronization of input modes during multimodal human-computer interaction *Proc. of the ACM SIGCHI Conf. on Human Factors in Computing Systems CHI '97 (Atlanta)* (New York: ACM) pp 415-22

[27] Maglio P P, Matlock T, Campbell C S, Zhai S and Smith B A 2000 Gaze and speech in attentive user interfaces In *Advances in Multimodal Interfaces—ICMI 2000* vol 1948, ed T Tan et al (Berlin Heidelberg: Springer) pp 1-7

[28] Prabhu V and Prasad G 2011 Designing a virtual keyboard with multi-modal access for people with disabilities *2011 World Congress on Information and Communication Technologies (WICT) (Mumbai)* (USA: IEEE) pp 1133-8

[29] MacKenzie I S and Zhang X 2008 Eye typing using word and letter prediction and a fixation algorithm *Proc. of the 2008 Symp. on Eye Tracking Research & Applications ETRA '08 (Savannah)* (New York: ACM) pp 55-8

[30] Soukoreff R W and MacKenzie I S 2001 Measuring errors in text entry tasks: an application of the Levenshtein string distance statistic *CHI '01 Extended Abstracts on Human Factors in Computing Systems (Seattle)* (New York: ACM) pp 319-20

[31] MacKenzie I S 2002 KSPC (keystrokes per character) as a characteristic of text entry techniques *Human Computer Interaction with Mobile Devices* vol 2411, ed F Paternò (Berlin Heidelberg: Springer) pp 195-210

[32] Soukoreff R W and MacKenzie I S 2003 Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems CHI '03 (Florida)* (New York: ACM) pp 113-20

[33] Ware C and Mikaelian H H 1987 An evaluation of an eye tracker as a device for computer input2 *Proc. of the SIGCHI/GI Conf. on Human Factors in Computing Systems and Graphics Interface CHI '87 (Toronto)* (New York: ACM) pp 183-8

[34] Porta M and Turina M 2008 Eye-S: a full-screen input modality for pure eye-based communication *Proc. of the 2008 Symp. on Eye Tracking Research & Applications ETRA '08 (Savannah)* (New York: ACM) pp 27-34

[35] Urbina M H and Huckauf A 2010 Alternatives to single character entry and dwell time selection on eye typing *Proc. of the 2010 Symp. on Eye Tracking Research & Applications ETRA '10 (Austin)* (New York: ACM) pp 315-22

[36] Helmert J R, Pannasch S and Velichkovsky B M 2008 Influences of dwell time and cursor control on the performance in gaze driven typing *Journal of Eye Movement Research* **2** 1-8

[37] Zhao X A, Guestrin E D, Sayenko D, Simpson T, Gauthier M and Popovic M R 2012 Typing with eye-gaze and tooth-clicks *Proc. of the 2012 Symp. on Eye Tracking Research & Applications ETRA '12 (Santa Barbara)* (New York: ACM) pp 341-4

[38] The View Point EyeTracker®: http://www.arringtonresearch.com/

[39] AbleData : http://www.abledata.com/abledata.cfm?pageid=113582&orgid=109986 , Accessed on Feb 11,2014

[40] Soft Switch : http://www.abledata.com/abledata.cfm?pageid=113583&top=0&productid=96746&trail=0, Accessed on Feb 11,2014

[41] USB Switch Interface : http://www.abledata.com/abledata.cfm?pageid=19327&top=15584&ksectionid=0&productid=75667&trail=22,11114,11131&discontinued=0 , Accessed on Feb 11,2014

[42] Abbott W W and Faisal A A 2012 Ultra-low-cost 3D gaze estimation: an intuitive high information throughput compliment to direct brain–machine interfaces *J. Neural Eng.* **9** 046016

[43] Morimoto C H and Amir A 2010 Context switching for fast key selection in text entry applications *Proc. of the 2010 Symp. on Eye Tracking Research & Applications ETRA '10 (Austin)* (New York: ACM) pp 271-4

[44] Vertanen K and MacKay D J C 2010 Speech dasher: fast writing using speech and gaze *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems CHI '10 (Atlanta)* (New York: ACM) pp 595-8

[45] Beelders T R and Blignaut P J 2012 Measuring the performance of gaze and speech for text input *Proc. of the 2012 Symp. on Eye Tracking Research & Applications ETRA '12 (Santa Barbara)* (New York: ACM) pp 337-40

[46] Hoste L, Dumas B and Signer B 2012 SpeeG: a multimodal speech-and gesture-based text input solution *Proc. of the Int. Working Conf. on Advanced Visual Interfaces AVI '12 (Capri Island, Italy)* (New York: ACM) pp 156-63

[47] Kumar A, Paek T and Lee B 2012 Voice typing: a new speech interaction model for dictation on touchscreen devices *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems CHI '12 (Austin)* (New York: ACM) pp 2277-86