

## Case Study

# A case study of developing gaze controlled interface for users with severe speech and motor impairment

D.V. Jeevithashree, Kamalpreet Singh Saluja and Pradipta Biswas\*  
*Indian Institute of Science, Bangalore, India*

### Abstract.

**BACKGROUND:** People with severe speech and motor impairment (SSMI) often depend on electronic user interfaces for communication, learning and many other daily activities. However, these interfaces are often designed assuming the preference and ease of use of end users for different screen regions is the same for people with SSMI as their able bodied counterparts. This paper presents a user study to evaluate whether users can undertake pointing and selection tasks faster if screen elements are organized at their preferred positions.

**OBJECTIVE:** To compare pointing and selection times in an eye gaze controlled interface between two conditions – screen elements randomly organized vs screen elements organized according to preference of users in terms of specific screen locations.

**METHODS:** We designed a word construction game using familiar 4-letter words and users were instructed to select the correct letters to construct words. We compared total times required to construct each correct word.

**RESULTS:** Users with SSMI can statistically significantly construct words faster [ $F(1,195) = 31.04, p < 0.01, \eta^2 = 0.14$ ] when letters were organized at their preferred screen positions than random organization.

**CONCLUSIONS:** Users with SSMI prefer middle and right side of screen more than the left side. Pointing and selection times in a gaze controlled interface can be significantly reduced by presenting screen elements at the preferred positions.

Keywords: Adaptive interface, eye gaze controlled system, assistive technology, SSMI, AAC

## 1. Background

This paper presents a case study of developing a gaze controlled interface for students with severe speech and motor impairment due to cerebral palsy. The case study followed a user centred design approach. Initially, we analysed visual search patterns of users and used these search patterns to design a user interface of a word construction game. A user study was undertaken to validate the user interface design. Finally, we have presented a gaze controlled Alternative and Augmentative Communication (AAC) Aid with an intelligent user interface that adapts positions of screen

elements based on frequency of use and ease of selection using eye gaze.

Eye tracking is the process of measuring either the point of gaze (where one is looking) or the motion of an eye relative to the head. An eye tracker is a device for measuring eye positions and eye movement. Most commonly used non-invasive eye gaze trackers are attached below a display and use pupil centre and corneal reflection technique [10]. Biswas and Langdon [6] reported a detailed literature survey on state-of-the-art gaze controlled interfaces and it may be noted that gaze controlled interfaces require either bigger button size and arrangement [17,18] or automatic zooming feature [2] or coupling with another interaction device [28] to accommodate inaccuracy in gaze tracking and micro-saccadic gaze movements.

\*Corresponding author: Pradipta Biswas, Indian Institute of Science, Bangalore 560012, India. E-mail: pradipta@iisc.ac.in.

### 1.1. State of the art

Most research in children with cerebral palsy is concentrated on developing applications like augmentative and alternative communication aids, adaptive menu structures [17,18], home automation applications [8] and so on. As representative of existing personalized AAC systems, the AVANTI [23] project addressed interaction requirements of individuals with motor disabilities and blindness using web-based multimedia applications and services. The AVANTI user interface could dynamically tailor itself to the abilities, skills, requirements and preferences of the users, to the different contexts of use, as well as to the changing characteristics of users, as they interact with the system. The CHAT [1] software proposed a predictive conversation model to achieve higher communication rate during conversation. This software predicted different sentences depending on situation and mood of the user. The user was free to change the situation or mood with a few keystrokes. Stephanidis and colleagues [24] presented a rigorous discussion on special HCI aspects for quadriplegic people. The Autonomia system replaced traditional windows and frame interface by a special interface designed to be operated by a single switch scanning technique. The Companions project [19] proposed to use telegraphic message as input and automatically produced grammatically correct sentences as output based on NLP techniques. The Friend project [4] used natural language generation techniques to construct grammatically correct sentences by taking a set of keywords from users. The KOMBE Project [20] tried to enhance communication rate by predicting a sentence or a set of sentences by taking sequence of words from users. The system was developed to cater Amyotrophic Lateral Sclerosis (ALS) patients. Yang et al. [27] proposed to use Morse code for an adaptable communication aid for users with physical disability.

There is already a plethora of commercial products [9,11] available for electronic gaze controlled interfaces for users with SSMI and researchers [16] already reported that gaze controlled interfaces provide “*new opportunities to communicate, interact and perform activities independently, as long as conditions are right*” while Borgsteig et al. [7] identified need for practice for long duration. However, it may be noted that none of existing AAC systems evaluated whether users have any preference for specific positions of elements on the screen and conducted any analysis on visual search patterns. They also do not adapt user in-

terfaces based on eye gaze fixation patterns of users. However, for mouse or other pointing devices, the Supple project [15] proposed to automatically adapt screen elements while the inclusive user model [3,5] simulated users’ interaction patterns and proposed personalizing interface based on simulation.

### 1.2. Our end users

In this particular study, our end users were all school students, quadriplegic and keen to learn operating a computer. The participants were secondary students at the spastic society of India in Chennai. All trials and interactions with them were undertaken under observation by their care takers and school instructors. All necessary permissions were taken before undertaking user trials. We took help from their teachers, who are rehabilitation experts, to evaluate their physical conditions. According to Gross Motor Function Classification system (GMFCS), they were all at level 5 as they could not move without wheelchair. According to the Manual Ability Classification System (MACS), some of them were at level 4 and rest were at level 5. A few of them could manage to move their hand to point to a non-electronic communication chart and others only relied on eye pointing. According to the Communication Function Classification System (CFCFS), all of them were at level 5 as they could not speak, could make only non-speech sound and communicate only through non-electronic communication board. They did not have access to any commercially available scanning software. Initially, we tried to use a mouse, joystick, trackball and stylus, but they could not manage to undertake any pointing and selection task using any of those devices as they could not make any precise movement using their hands necessary to control those devices. Their teachers and parents informed us that they were accustomed to use eye pointing with non-electronic communication chart. We have described more details on individual users in the following table. All of these users did not take part in all studies, we identified them by their codenames in subsequent studies.

### 1.3. Pilot studies

Before using or developing any eye gaze controlled software, we undertook a series of studies to investigate differences in fixation patterns and eye gaze movements of users with SSMI compared to their able-bodied counterparts. Researchers already investigated

Table 1  
Description of participants

Participant code	Age (years)	Gender	Education	Description
A	14	F	NIOS National Institute of Open Schooling Standard-10th	She understands logic easily. She has no finger isolation i.e. she cannot use hand to point something. She cannot press a switch with her hands because of tremors athetoid movement. She has tremor in head and chin movements. She can respond yes/no by nodding head. She is comfortable with eye tracking.
B	20	F	NIOS-12th standard	She takes time to understand things. She cannot hold things with her hand. She lacks finger isolation. She needs help of an assistant for keeping her head straight. She responds to yes/no by nodding her head. Her voice bottles up.
C	15	M	NIOS-10th standard	His hands are very rigid. He has rigid body and head movement. He responds by gentle blinks to say yes.
D	06	F	STATE Syllabus-3rd standard	She cannot speak but will respond to general talks by making voices. She is not so fluent in her hand movements but can use her hand to point something. She likes to interact with people by taking inputs from an assistant.
E	06	M	NIOS 2nd standard	He gets tensed easily. He is keen in listening to stories and loves to interact with people. He uses eyes.
F	12	F	Prevocational training	She is shy and tries to speak with a very low voice. She gets distracted easily. She has control of her head and hand movement. She always keeps her head down.
G	7	M	NIOS 2nd standard	He is a quick learner. He can communicate well with his eyes. He can point using his hands and has good finger isolation. He can sight words on a spell chart comfortably.
H	20	F	NIOS 12th standard	She finds her interest in multimedia and can access computers well. She can use her hand to hold things like spoon. Her hands possess athetoid movement. She can speak but with poor clarity. She can use spell chart and qwerty chart. She likes to interact with people. She has passion for making accessories. She can control head movement but has tremors.
I	23	M	NIOS 12th standard	He uses spectacles but is comfortable with eye tracking. He cannot speak properly. He responds to people by making sound and moving his head. He is scared of sudden sounds. He can ride his wheel chair himself.
J	22	M	NIOS 12th standard	He uses spell chart and gestures for communication. He cannot speak differential vocalization. He lacks isolation of fingers and cannot hold things with his hands. He can respond yes/no well by nodding head. He can ride wheel chair himself. He wishes to open an automobile shop/electrical shop. He gives exam using a qwerty chart.
K	15	F	NIOS 10th standard	She does not have any control over her body especially from hip and above region. She is good at doing any task given.

visual function in children with cerebral palsy and reported presence of nystagmus in less than 10% in one representative sample [26] and about 72% in another sample [12]. Nystagmus was also accompanied by loss of visual acuity, contrast sensitivity and strabismus. Fazzi et al. [12] reported that “*clinical expression of cerebral visual impairment can be variable*” requiring a case by case analysis of end users.

### 1.3.1. Study 1 – Analyzing fixation patterns

This study investigated whether our end users can fixate attention to a visual stimuli and the duration of their saccadic eye gaze movement before a gaze tracker detects a fixation near the stimuli. Previous works by Penkar [20] and Nayar [19] investigated optimizing dwell time with respect to target size, position and his-

tory of use but we wanted to first investigate the spatial distribution of eye gaze with respect to visual stimuli. We also undertook comparative analysis between users with SSMI and able bodied counterpart of the eye gaze locations recorded by the eye gaze tracker with respect to a visual stimulus.

**Participants:** We collected data from 12 participants – 6 participants were users with SSMI (A, B, C, D, G, H) while the rest were able bodied students (3 male, 3 female, age range 19 to 25 years).

**Design:** The study displayed a 5 mm × 5 mm white stimulus at a random position in a black background for 2 secs and then a blank screen for 2.5 secs. The process was repeated for 5 minutes for each participant.

**Procedure:** Initially participants went through the 9 points calibration procedure of the Tobii gaze tracker.

Then they were only instructed to fixate attention on the white stimuli as soon as it appeared on screen.

**Results:** Initially we investigated eye gaze positions of participants while the visual stimulus was shown on screen. We calculated the offset (difference) of the position of recorded gaze positions and the stimulus while it was visible on screen. Figure 1 below plots the histograms of x and y deviations for both user groups – blue bars represent users with SSMI while orange bars represent their able-bodied counterpart.

Next, we investigated the minimum time required to record eye gaze position within 50 pixels of the stimulus. Instead of the raw eye gaze position, we compared performance of an averaging and median filter. The filter runs a sampling window of latest 10 gaze locations and return either the arithmetic mean or median of the latest 10 eye gaze position. For both groups, the peak occurred at 400 msecs for median filter and at 450 msecs for averaging filter. We also noted that within 1500 msecs, we could record a gaze position near the stimulus for 82% of cases for users with SSMI and 98% cases for their able-bodied counterpart.

It may be noted that the peak occurs between  $-50$  and  $+50$  pixels indicating both group could fix attention on visual stimulus and the eye gaze tracker was able to record it. However, the standard deviation of offsets was almost 2 times higher for the spastic group than their able-bodied counterpart. We also calculated correlation between the distance of the stimuli from center of screen and mean offset and angular deviation of the visual stimuli and mean offset but the coefficients were less than 0.3 for both groups of users.

A cumulative histogram of the offsets (Fig. 2) shows that if the target size is more than 50 pixels ( $1.13^\circ \times 1.10^\circ$  of visual angle) then the probability of selecting a target through a saccade within the target area increases by 25% and if the target is more than 100 pixels then the probability increases by 35% and so on.

**Discussion:** This study aims to investigate response to visual stimuli by users with SSMI with an aim to calibrate gaze controlled cursor control device. The study demonstrates that users with SSMI could fixate attention although have more uncontrolled saccadic gaze movements than their able-bodied counterparts. The offset did not correlate with screen position or angular deviation of the stimuli. The size of target can be optimized by analyzing the offsets. We also compared a median and mean filter to reduce effect of uncontrolled gaze movements and noted that both able-bodied and spastic participants can fixate attention to visual stimuli within 1.5 secs in more than 80% cases. The me-

dian filter found to response 50 msecs faster in detecting fixation compared to the averaging filter.

In subsequent studies, we calculated median from recorded gaze points in a 400 msecs time window and the screen pointer was moved based on the value of the median. Selection was performed by dwelling the pointer for 1500 msecs.

### 1.3.2. Study 2 – Analysing visual search pattern

This study aimed to compare visual search patterns between users with SSMI and their able-bodied counterpart. Unlike Feit [13]'s study, we did not measure precision and accuracy for different screen regions, rather we used a nearest neighborhood algorithm, that activates target nearest to the cursor location and with that we measured users' preference and performance for different screen regions.

**Participants:** We collected data from 20 users – all 11 users with SSMI and nine able-bodied students (6 males, 3 females, age range 19 to 25 years).

**Design:** We displayed a set of ten balloons (Fig. 3) on the screen. Each balloon was  $103 \times 112$  pixels in size. Participants were requested to point and click on all balloons. The balloons disappeared on clicking. When all balloons disappeared, a new set of ten balloons appeared on screen. We did not specify any particular order to select balloons.

We implemented the following algorithm to point and click on the balloons using eye gaze. We calculated the median of gaze position every 400 msecs. The cursor moved on the screen based on the median gaze position. The balloon nearest to the position of the cursor was enlarged to 1.5 times its size. If the gaze dwell near or on the balloon for 1.5 secs, it was selected and disappeared from screen.

**Procedure:** Initially, participants undertook the 9 points calibration routine provided by Tobii SDK [24]. Then they undertook a training session and after they understood the task, they were instructed to point and click all balloons. We recorded at least 15 pointing and clicking tasks from each participant.

**Results:** We have investigated the following four dependent variables:

1. *Pointing and selection times for each position:* Total time spent between a selection and the next one
2. *Frequency of first choice:* Which position was first selected and how many times
3. *Frequency of selection:* How many times each position was selected
4. *Patterns:* Sequences of selections

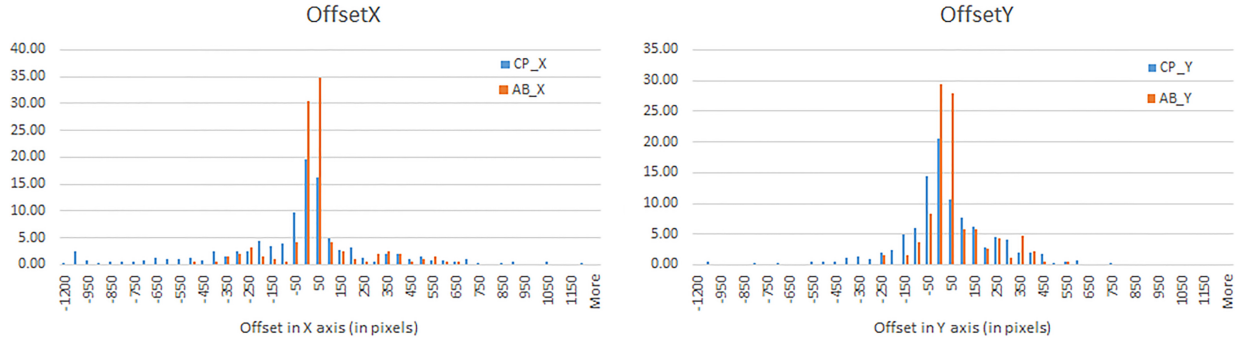


Fig. 1. Comparing eye gaze positions with respect to visual stimuli.

Cumulative Histogram of Offset for users with CP

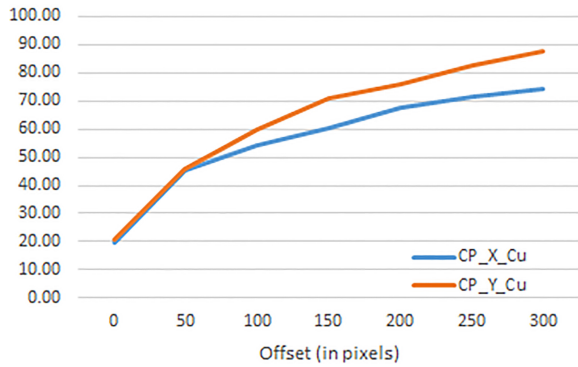


Fig. 2. Cumulative histogram of offset.

For each dependent variable, we compared performances of users with SSMI with their able-bodied counterparts. Pointing and selection times were significantly lower for able-bodied users compared to users with SSMI [ $t(1,9) = -12.95, p < 0.001$ ]. The variance was 38 times higher for users with SSMI indicating they occasionally took long time to make a selection.

We analyzed the pointing and selection times for each position of targets for both user groups and Fig. 4 indicates lower selection times with bigger font size. It may be noted that users with SSMI took least time to select balloons at Top Right (TR), Middle Centre Right (MCR) and Bottom Middle (BM) positions while the able-bodied group took least time for Bottom Left (BL), Middle Centre Left (MCL) and Middle Right (MR) positions.

Figure 4 indicates the frequency of selection of first position by drawing a black border around the three most frequent positions. Users with SSMI most of the time first selected one of the MCL, MCR and TR buttons while their able-bodied counterparts most of the time first selected Top Left (TL), Top Middle (TM) and MCR buttons.

Although participants were instructed to select all 10 balloons but users with SSMI often could not select all buttons in the screen. The standard deviation among the number of selections summed up for each individual position is only 1 for able-bodied users while it is 8.8 for users with SSMI. Figure 4 indicates the total number of selections according to position using color coding. The green color indicates the first three preferences, yellow next four and red indicates the least three. The number of selections and average pointing and selection times with respect to each position was negatively correlated ( $r = -0.29$ ) for users with SSMI and positively correlated ( $r = 0.43$ ) for able-bodied users. None of these correlations were significant at  $p < 0.5$ .

Finally, we analyzed the patterns in sequences of selections – means we investigated how many times users select button A after button B for all pairs of values of A and B and similarly all possible patterns of selections consisting three consecutive selections. Figure 4 shows the top most 2-buttons and 3-buttons sequences using blue and brown arrows respectively. We only marked the sequences which appeared at least four times or more and the thickness of the arrow indicates the frequencies of occurrences of the patterns. It may be noted that users with SSMI had MCR-TM, TL-ML-TM as two frequent sequences of gaze movements occurring more than 4 times which included right to left and bottom to top movements while for able-bodied users all frequent sequences were from left to right and top to bottom.

**Discussion:** The study shows users with SSMI took longer with more variance to point and select than their able-bodied counterparts. We also noted a left to right and top to bottom search strategy for able-bodied users while the frequency of total selection, first selection and visual search patterns indicate a nearest neighborhood strategy for users with SSMI. The nearest neighborhood strategy for users with SSMI.

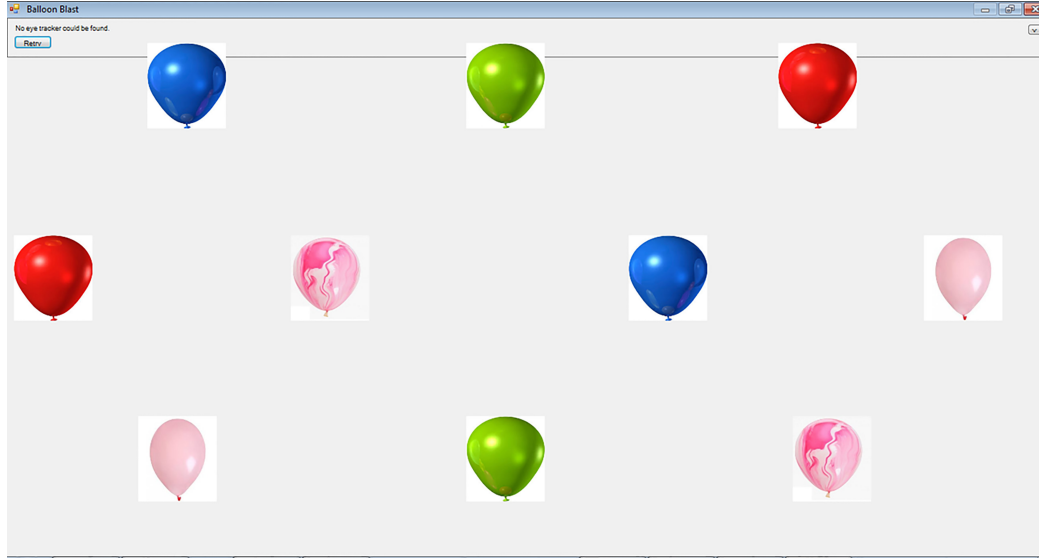


Fig. 3. Design of task for analyzing visual search pattern.

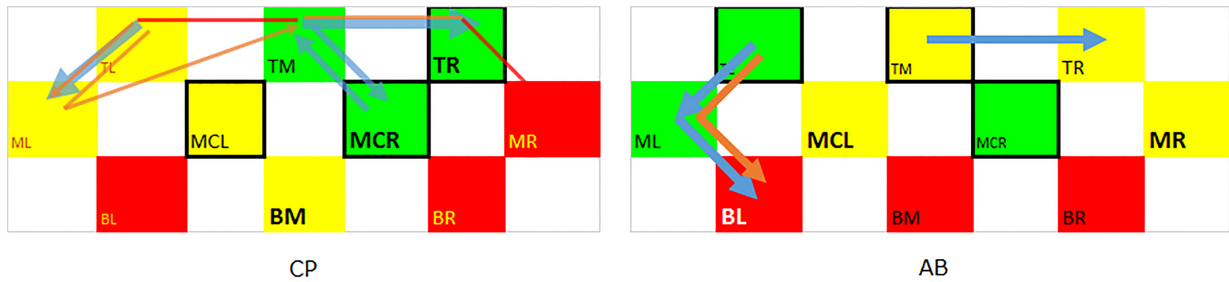


Fig. 4. Comparing patterns of selections with respect to screen positions between users with SSMI and their able-bodied counterparts.

neighborhood strategy means users selected the nearest target from their present position instead of going through a serial scanning technique. The task initiated with the focus at the middle of the screen and then it can be noted that MCR, TR and TM positions were mostly selected and reaction time was also lowest for MCR, TR and BM positions. The highest observed patterns also indicated this nearest neighborhood strategy instead of left to right and top to bottom search strategy. This search strategy could be leveraged while developing software interface for users with SSMI, earlier Fleetwood [14] reported similar strategy in an icon searching task even for able-bodied users. However, the pointing and selection times were not related to the number of selections. It may indicate users' search strategy is independent of the time they require to select a target.

However, in this study we could not test statistical significance with respect to positions of targets. The following study investigates only two different screen organization with respect to target positions and we

undertook statistical hypothesis testing for users with SSMI and able-bodied users separately.

## 2. Proposed approach

We initially developed a software to control a cursor using eye gaze and this software can be used to operate any application of a MS Windows operating system. Then we developed a simple application to evaluate whether users' pointing and selection times improve if we place screen elements supporting their eye gaze fixation and search strategy. Finally, we proposed an AAC system with an adaptable interface supporting their eye gaze fixation and movement strategy.

### 2.1. Gaze controlled cursor

We developed the following algorithm for controlling an on-screen cursor using a screen-mounted eye

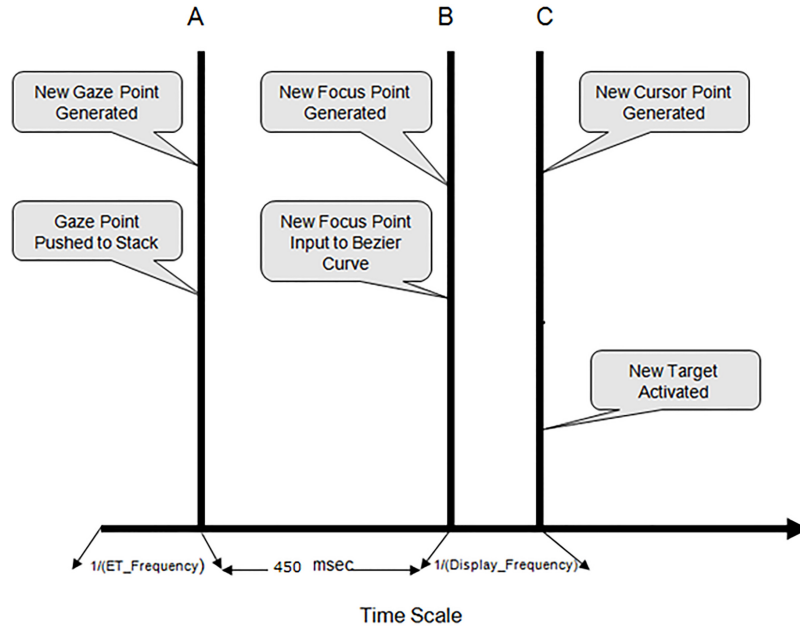


Fig. 5. Gaze controlled cursor movement algorithm.

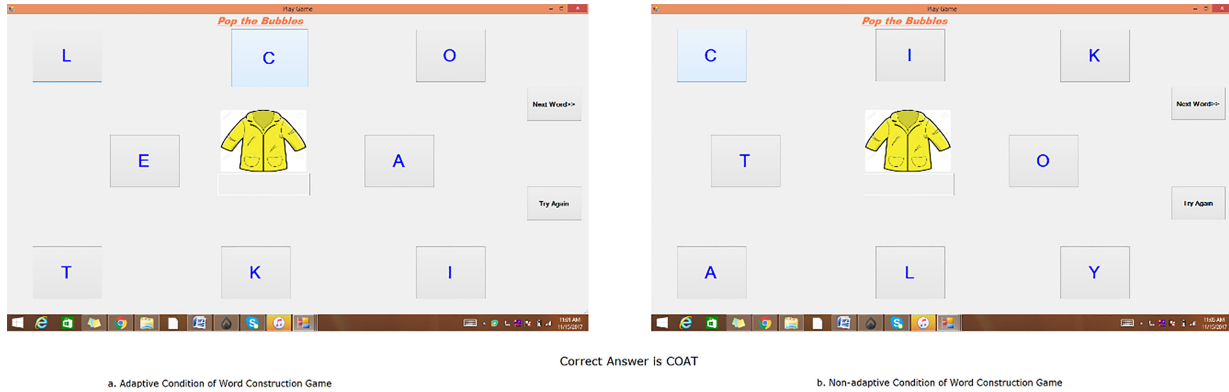


Fig. 6. Gaze controlled word construction software.

gaze tracker. Our gaze tracking system records the eye gaze positions continuously (refer point A in Fig. 5) and takes the median of the pixel locations in every 450 msec to estimate the region of interest or saccadic focus points (refer point B in Fig. 5). The median was less susceptible to outliers than the arithmetic mean in case the eye gaze tracker briefly lost signal or in case of nystagmus of users. We simulated the eye movement using a Bezier curve that smoothens the cursor movement between two focus points as we wanted to make the cursor movement looking similar to existing cursor control devices like mouse or trackball. The algorithm pushes the focus points into a queue data structure and the Bezier curve [21] algorithm interpolates points in

between two focus points (refer point B in Fig. 5). The pointer is drawn at each interpolated points in every 16 msec to visualize a smooth on-screen movement (refer point C in Fig. 5). Based on the present cursor position, we activated an on-screen element or target.

## 2.2. Gaze controlled interface

We have designed a simple word construction game to be operated by selecting individual letters. Figure 6 below shows screenshots of the interfaces. The participant was instructed to select letters to construct a word describing the picture in the middle. All words were 4-letters words and each screen had only one correct

answer. The font sizes were bigger than 14 pt and all end users had either 6/6 visual acuity or used corrective glasses.

We used this interface to evaluate preference on screen positions for users with SSML. In one version of the software, hereafter referred to as non-adaptive condition, we randomly placed all letters on the screen based on a uniform distribution. In another condition, hereafter referred to as adaptive condition, we placed the correct letters in preferable positions using a normal distribution. We placed each letter using a function that generates a random number from a normal distribution. If the random number is within the first standard deviation, we placed the correct letter in the preferred slot, otherwise it was placed in a non-preferred slot. We also implemented a nearest neighbourhood predictor for the adaptive condition. Using this algorithm, the participant can select a target even when the pointer is not on the target button but only near the desired target. A video demonstration of the software can be found at <https://youtu.be/UOLG6QkZ3mI>. The following section describes a user study comparing these two interfaces.

### 2.3. Validation study

This study evaluated whether placing screen elements at preferred position can reduce pointing and selection times for users with SSML.

**Participants:** We collected data from 12 participants – 6 of them were users (age range 12 to 19 years, 2 males, 4 females) with cerebral palsy (A, B, C, D, H, I) and 6 were there able-bodied counterparts.

**Material:** We used an Intel NUC computer running Windows 7 operating system and a 15" display for displaying the stimulus. Eye gaze was recorded using a Tobii PCEye mini eye gaze tracker. The cursor was moved using eye gaze following the algorithm presented in the previous section of this paper.

**Design:** Participants constructed 5 words using both the adaptive and the non-adaptive versions of the word construction software using the gaze controlled cursor movement algorithm presented in the previous section. We undertook a  $2 \times 2$  ANOVA with the following factors

1. User
  - a. Users with cerebral palsy
  - b. Control group
2. Software
  - a. Adaptive

#### b. Non-adaptive

The order of the conditions was altered between each pair of participants; half of them undertook the trial using the adaptive condition and half using the non-adaptive condition. We recorded the timestamp of each selection.

**Procedure:** Initially, participants were briefed about the purpose of the study. Then they undertook the 9-points calibration procedure for the Tobii tracker. They went through a training session and finally undertook the actual trial.

**Results:** We investigated the differences between the instances of two consecutive button selections. This time, hereafter referred to as button selection time, consists of the visual search time for the target button and the pointing and selection time using the gaze controlled system. Initially, we screened data for outliers and removed seven samples based on the values of outer fence ( $Q3 + 3 \times (Q3-Q1)$ ). We investigated average button selection times for each individual participant and we noted that all users with cerebral palsy required less time on average to select buttons in the adaptive condition [(mean = 8.9 secs; SD = 6.5 secs) compared to the non-adaptive condition (mean = 16.5 secs; SD = 15.3 secs)]. Similarly, the control group required less time on average to select buttons in the adaptive condition [(mean = 5.6 secs; SD = 3.3 secs) compared to the non-adaptive condition (mean = 5.9 secs; SD = 3.8 secs)]. In each of adaptive and non-adaptive conditions, we recorded more than 100 button selection tasks. We found

- A main effect of user [ $F(1,195) = 150.71, p < 0.01, \eta^2 = 0.44$ ]
  - A main effect of software [ $F(1,195) = 31.04, p < 0.01, \eta^2 = 0.14$ ] and
  - An interaction effect of user and software [ $F(1,195) = 26.25, p < 0.01, \eta^2 = 0.12$ ]
- Then we undertook a couple of pairwise unequal variance t-tests and noted that
- For users with cerebral palsy there is a significance difference in button selection times between adapted and non-adapted conditions [ $t(1, 150) = -4.72, p < 0.01$ ].
  - For the control group, the difference in button selection times between adapted and non-adapted condition was not significant at  $p = 0.05$ .

In Fig. 7 below, we plotted two box plots for button selection times (in msec) for both user groups. It may be noted that for users with cerebral palsy, both average and standard deviation of button selection times were



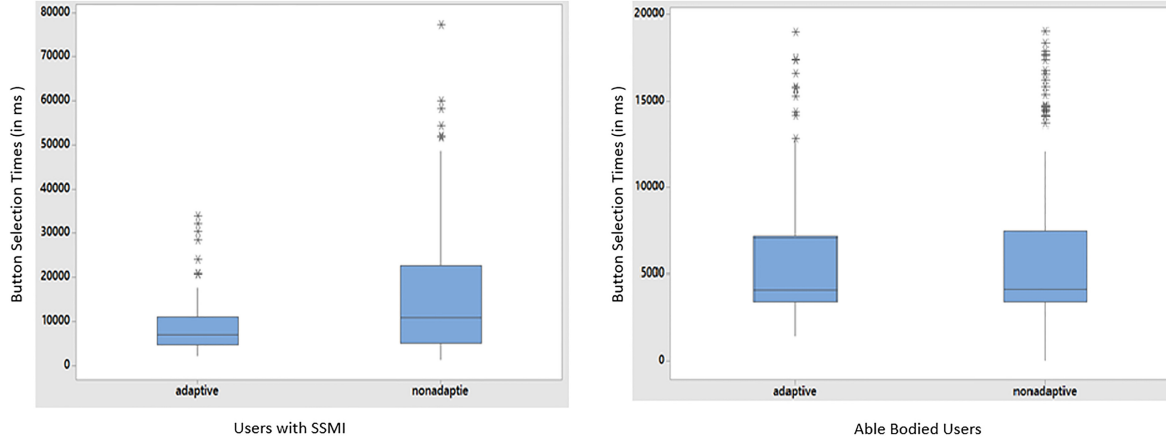


Fig. 7. Box plots of button selection times.

reduced in the adaptive condition compared to the non-adaptive one.

Finally, we investigated average button selection times for each individual participants and we noted that all users with cerebral palsy required less time on average to select buttons in the adaptive condition while 4 out of 6 users in the control group took less time in the adaptive condition than the non-adaptive one. In Fig. 8,  $P_i$  stands for  $i^{\text{th}}$  user with cerebral palsy while  $C_i$  stands for the control group.

**Discussion:** The user study demonstrates that a graphical user interface designed according to the visual search strategy can significantly improve user interaction by reducing visual search and pointing and selection times in a gaze controlled interface. It may be noted here that the purpose of the software or study was not to design a game – we assumed that the correct letters were more likely to be selected than incorrect ones and placing the correct ones in favorable positions will reduce visual search and pointing times. The same principle can be followed in developing AAC systems with letter or word prediction features by placing the more common words or letters in middle or middle-right positions on the screen.

However, we noted that the average selection time (which also includes visual search time) was still about 6 secs even in the adaptive condition for users with Cerebral Palsy. Our user group used this particular software and gaze controlled system for the first time during this trial.

### 3. Application on developing an AAC system

We have developed an assistive communication board software for children having spasticity which re-

arranges screen elements based on preferred position of users as described in the previous studies. Presently our end users require an intermediary to translate communication from a printed communication board to meaningful sentences. AAC (Augmentative and alternative communication) can help these children express themselves and connect with family, caregivers, and others. This software presents an eye gaze controlled AAC platform which reduces the need of a third person by allowing the children to operate the pictorial communication board having text to speech methods. In particular, the user may gaze at a picture to articulate a respective phrase using the TTS (Text-To-Speech) feature of the software. Based on the inputs of our studies described in previous sections, we placed screen elements in the centre and right positions of the screen as our end users found it difficult to move their eye gaze towards the left side of the screen. Buttons were arranged in a fashion such that the most frequently used buttons appeared on the first page for quick access. We used a nearest neighbourhood predictor algorithm to identify the nearest screen element to the current cursor position making it easier for users to select the target button.

**Buttons rearrangement:** The rearrangement of buttons was done based on recency and frequency of use. Recency measures how recently an on-screen element is selected and frequency is measured by how many times a particular element is selected. We have considered the fact that weight of recency should be more than frequency. We have computed a weighted score from timestamp and frequency and sorted the list of screen elements accordingly. The algorithm tries to put frequently and recently used elements on the middle and right side of the screen and remaining elements

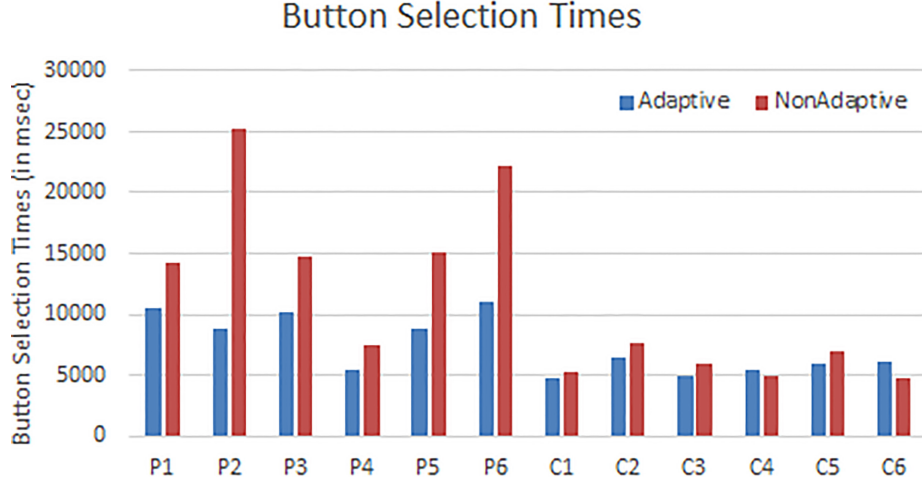


Fig. 8. Average button selection times for each participant.

to the left side, but not on extreme left. The weighted score is computed using the following formula:

$$weighted\_score(T_i, F_i) = \frac{N_t}{N_f + N_t}$$

In the above equation,  $N_t$  is normalised timestamp and  $N_t = T_i/T$  where  $T_i$  and  $T$  are the current timestamp and mean of all the timestamps respectively.

$N_f$  is normalised frequency and  $N_f = F_i/F$  where  $F_i$  and  $F$  are the current frequency and mean of all the frequencies respectively.

We have computed weighted scores for every on-screen element and passed it to following sorting function:

```

begin Sort(list_of_single_factors)
list = list_of_single_factors
for all elements of list
    if list[i] < list[i+1]
        swap(list[i], list[i+1])
    end if
end for
return list
end Sort

```

The pseudocode above generates a list of screen elements sorted in descending order of recency and frequency of use. The following formulae gives us page number,

$$page\_number(index) = \left\lceil \frac{index}{Maximum\ number\ of\ buttons\ displayed\ for\ each\ page} \right\rceil$$

Where,  $\lceil x \rceil = \text{ceiling}(x)$ , ceiling function maps  $x$  to the least integer greater than or equal to  $x$ , Index = index of the picture button in the sorted list.

The following steps determine the position of different elements on the screen:

Coordinates for the centre element are calculated as follows:

Step I :

$$X = \frac{W_s - W_b}{2}$$

$$Y = \frac{H_s - H_b}{2}$$

Where  $W_s$  and  $W_b$  are the screen width and button width respectively, and  $H_s$  and  $H_b$  are the screen height and button height respectively.

We considered the screen as a circle with centre coordinates and divided it into two halves vertically. The radius of the circle depends on the value of  $N$  (total number of screen elements) and screen dimensions. We first populated the right side of the screen with recently and frequently used elements and then arranged rest on the left semicircle.

For the right semicircle,

Step II:

Calculating, the position of  $X$  and  $Y$  coordinates for the rest of the buttons,

$$X = \frac{W_s}{2} + r * \cos(\alpha),$$

$$Y = \frac{H_s}{2} + r * \sin(\alpha)$$

Where,

$W_s$  is the available screen width,

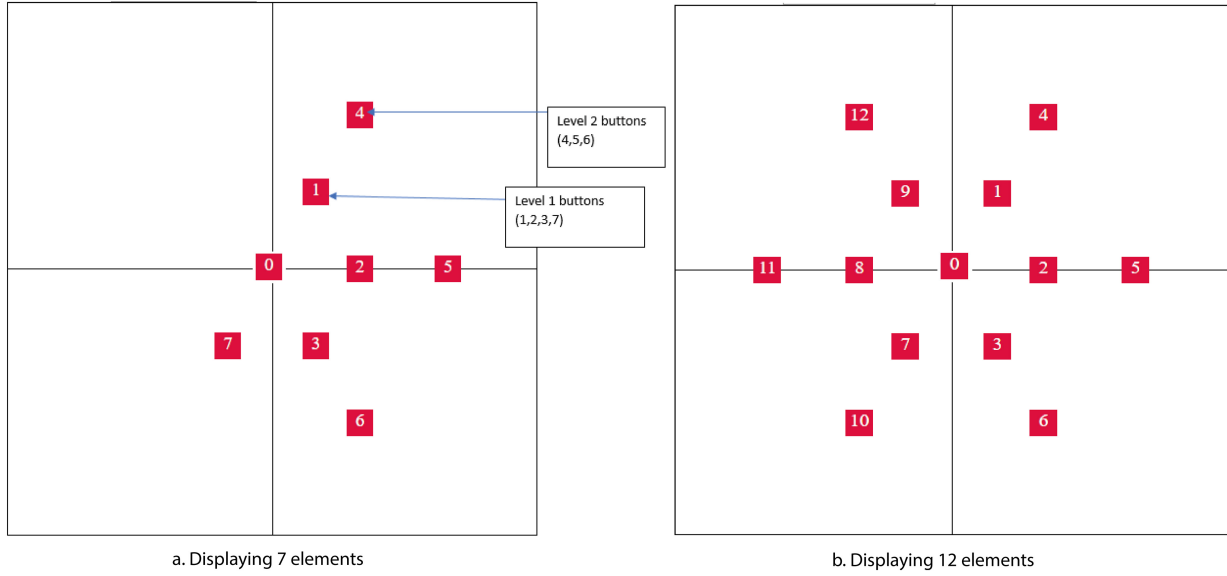


Fig. 9. Explaining output of algorithm in terms of screen elements placements.

$H_s$  is the available screen height,

$r$  is the radius of the circle,

$\alpha$  is the cumulative fixed\_angle calculated as

$$\text{fixed\_angle} = \frac{2\pi}{N}$$

Where  $N$  is the total number of buttons to be displayed on the screen.

The algorithm repeats step II for the left semicircle. The calculated coordinates were adjusted to lie inside the screen viewport. If the algorithm could not place all buttons after step II, it followed step III for the right semi-circle

Step III:

For the next level of buttons,

$$r = l * r,$$

where,  $l$  is the number of level,  $l = 1, 2, 3, \dots, n$ ,  $r$  is the radius.

The algorithm repeated step III for the left semicircle if it could not place all buttons in the right semicircle. Figure 9 above shows a sample output from the algorithm, the red boxes indicating screen elements and the number inside indicating their recency and frequency of use based on the formula described above.

Besides the button rearrangement feature, the proposed AAC application has the following features:

- 1) Each screen element consists of a picture, caption, phrase, and a voice button.
- 2) The next and previous buttons provide functionalities for going to next and previous pages respectively.

3) Selecting an element with eye gaze articulates its respective phrase.

4) The software has male and female voices. The phrase can be articulated in any of the male or female voices.

5) We provided a facility for the instructor to change the picture, caption or phrase in the picture button by right-clicking on it.

6) The instructor also has the facility to add new or delete existing screen elements.

Figure 10 shows a sample screenshot of the proposed system.

In Fig. 10, the maximum number of buttons displayed for each page without occlusion is 6 with respect to screen dimensions. The picture button with higher value of weighted score is placed at the centre on each page and the rest of picture buttons are placed radially, first on right side and then on left. No elements have been placed at the extreme left side of the screen.

To test the button rearrangement feature, we installed the software at the spastic society on a computer running Windows 7 operating system and attached to a Tobii PCEye mini eye gaze tracker. Five students with SSMI used the system. All of them also took part in the studies described earlier in the paper. Each of them used the system for approximately 5 to 10 minutes. We did not yet run a controlled experiment with the software, rather recorded only the adaptation feature. We noted that when a participant

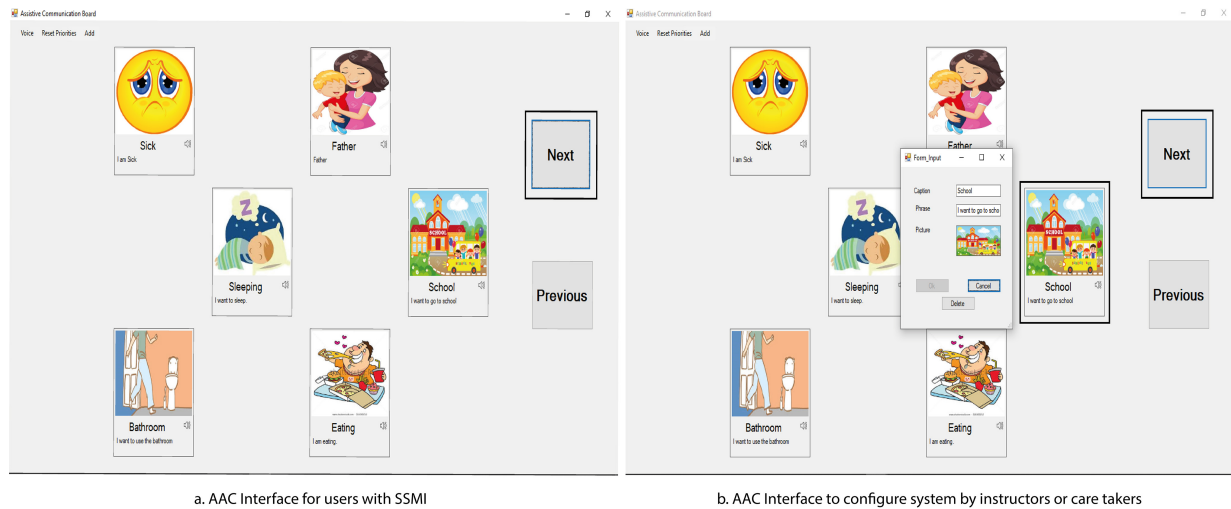


Fig. 10. Proposed AAC system.

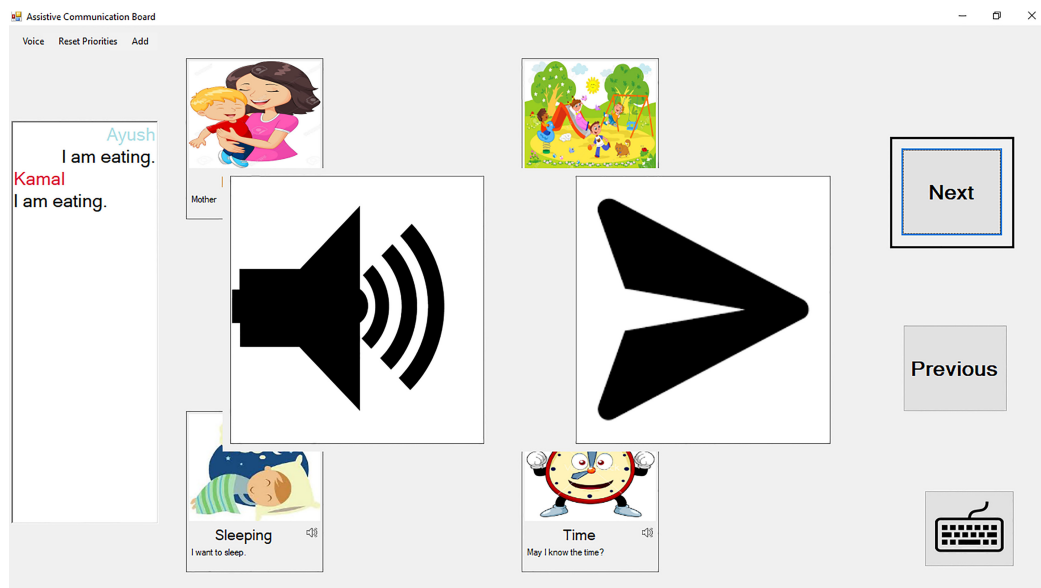


Fig. 11. Using the AAC system as a chat application.

- clicked button with caption “Sleep” for 5 times, it moved from bottom left on page number 1 to centre on page number 1.
- clicked button with caption “Father” for 10 times, it moved from centre left in the page number 2 to centre right on page number 1.
- clicked button with caption “Food” for 7 times, it moved from the top left on page 1 to top right on page number 1.

Overall, all five participants were able to use the system and we left the set up with the software at the spastic society for longer term use.

The same application can also be used to send messages to another participant. A participant has two options upon clicking on a picture – speak button (left) or send button (right) (Fig. 11). The application also has a virtual keyboard interface which supports word auto completion feature. A participant can switch between keyboard interface and picture-based communication interface by gazing on the button which is placed on the bottom right corner of screen.

We recorded two scenarios of interaction by four users as described below.

*Scenario1:* Participants A and B used the application together for about 14 minutes. Participant A and B started with the picture-based communication interface. Participant A selected the picture “Hello” and sent it to participant B. It took 5 seconds to select and send the word. Participant B received the message on the left panel and replied with the picture “Hi”. It took 7 seconds for Participant B to gaze at the message and send to Participant A. Afterwards, they exchanged following messages – How are you, Good, I want to go home, No, I like pizza, Good, I want to go to Class,. Then, Participant B selected the virtual keyboard interface. Participant B typed the message “My name is ABC”. It took 4 minutes 53 seconds to type and send the message. Participant A received the message and replied with picture “Nice”. It took 12 seconds for Participant A to send the message to Participant B. Then, Participant A selected the virtual keyboard interface. Participant A typed the word “My name is XYZ”. It took 6 minutes 36 seconds to type and send the message.

*Scenario2:* Participants C and D used the application together for 15 minutes. Participant C and D started with the virtual keyboard interface. They spent 12 minutes on virtual keyboard interface. They exchanged following messages – XYZ is good friend, My name is DEF. The actual names were long requiring longer duration. They spent 3 minutes on picture-based communication interface. They exchanged following messages I am good, Yes, How are you? Do you want to play? Very good, I am hungry, I am eating, Stop, OK, Goodbye.

#### 4. Conclusions

This paper presents a case study through a user centred design approach for developing an gaze controlled Augmentative and Alternative Communication Aid for users with severe speech and motor impairment (SSMI). We undertook a series of studies on eye gaze fixation and movement patterns of users with SSMI and noted they prefer the middle and right side of the screen for fixating attention more often than the extreme top and left side. We described a study involving a gaze controlled word construction game that showed users can undertake pointing and selection tasks statistically significantly faster if the screen elements are placed at their preferred positions. Finally, we proposed an AAC system and chat application that places screen elements in preferred positions and also adapts positions of screen elements based on their recency and frequency of use.

#### Conflict of interest

None to report.

#### References

- [1] Alm N, Arnott JL, Newell AF. Prediction and conversational momentum in an augmentative communication system, *Communications of the ACM* May 1992; 55(5).
- [2] Bates R. Multimodal Eye-Based Interaction for Zoomed Target Selection on a Standard Graphical User Interface. *Proceedings of INTERACT*, 1999.
- [3] Biswas P, Robinson P. Modelling User Interface for Special Needs, *Proceedings of the Accessible Design in the Digital World (ADDW)*.
- [4] Biswas P, Samanta D. Friend: a communication aid for persons with disabilities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 16(2), 205-209.
- [5] Biswas P, Robinson P. Evaluating the design of inclusive interfaces by simulation, *Proceedings of the 15th ACM International conference on Intelligent User Interfaces (IUI)*, pp. 277-280.
- [6] Biswas P, Langdon P. Multimodal Intelligent Eye-Gaze Tracking System. *International Journal of Human-Computer Interaction* 31(4), Taylor & Francis, Print ISSN: 1044-7318.
- [7] Borgestig M, et al., Eye gaze performance for children with severe physical impairments using gaze-based assistive technology – A longitudinal study. *Assistive Technology* 2016; 28(2): 93-102, <http://dx.doi.org/10.1080/10400435.2015.1092182>.
- [8] Castellina E, Razzak F, Corno F. Environmental Control Application compliant with Cogain Guidelines, *Proceedings of the 5th Conference on Communication by Gaze Interaction – Gaze Interaction for Those Who Want It Most COGAIN*, 2009.
- [9] Communicate with the world using the power of your eyes, Available at <http://www.eyegaze.com/eye-tracking-assistive-technology-device/>, Accessed on 8/11/2017.
- [10] Duchowski AT. *Eye Tracking Methodology*. Springer, 2007.
- [11] Eye-gaze control technology, Available at <https://www.cerebralpalsy.org.au/about-cerebral-palsy/interventions-and-therapies/eye-gaze-technology-for-children-and-adults-with-cerebral-palsy/#1473737787180-c2ceacf9-ff87>, Accessed on 8/11/2017.
- [12] Fazzi E, et al., Spectrum of Visual Disorders in Children with Cerebral Visual Impairment. *Journal of Child Neurology* 2007; 22(3): 294-301.
- [13] Feit AM, et al., Toward Everyday Gaze Input: Accuracy and Precision of Eye Tracking and Implications for Design, *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1118-1130.
- [14] Fleetwood MF, Byrne MD. Modeling the visual search of displays: a revised act-r model of icon search based on eye-tracking data. *Human-Computer Interaction* 2006; 21(2): 153-197.
- [15] Gajos KZ, Wobbrock JO, Weld DS. Automatically generating user interfaces adapted to users’ motor and vision capabilities. *ACM Symposium on User Interface Software and Technology* 2007; 231-240.
- [16] Holmqvist E, Thunberg G, Dahlstrand MP. Gaze-controlled communication technology for children with severe multiple disabilities: Parents and professionals’ perception of gains,

- obstacles, and prerequisites. *Assistive Technology* 2017, DOI: 10.1080/104004352017.1307882.
- [17] Majaranta P, Majaranta N, Daunys G, Špakov O. Text Editing by Gaze: Static vs. Dynamic Menus, Proceedings of the 5th; Conference on Communication by Gaze Interaction – Gaze Interaction For Those Who Want It Most COGAIN, 2009.
  - [18] Mario H, Urbina MH, Lorenz M, Huckauf A. Selecting with gaze controlled pie menus, Proceedings of the 5th Conference on Communication by Gaze Interaction – Gaze Interaction for Those Who Want It Most COGAIN, 2009.
  - [19] McCoy K, et al., Some Interface Issues in Developing Intelligent Communication Aid for Disabled. ACM International Conference on Intelligent User Interfaces 1997, Orlando, Florida USA.
  - [20] Pasero R, Richardet N, Sabatier P. Guided Sentences Composition for Disabled People, Proceedings of the fourth conference on Applied natural language processing, October 1994.
  - [21] Salomon D. *Curves and Surfaces for Computer Graphics*, Springer Verlag, August 2005, ISBN 0-387-24196-5.
  - [22] Stephanidis C, Paramythis A, Sfyarakis M, Stergiou A, Maou N, Leventis A, Paparoulis G, Karagiannidis C. Adaptable And Adaptive User Interfaces for Disabled Users in the AVANTI Project, *Intelligence In Services And Networks*, LNCS-1430, Springer-Verlag 1998, 153-166.
  - [23] Stephanidis C, et al., Designing human computer interfaces for quadriplegic people, *ACM Transactions on Computer-Human Interaction* June 2003; 10(2): 87-118.
  - [24] Tobii SDK. Available at <https://developer.tobii.com/>, Accessed on 24/01/2019.
  - [25] Tobii PCEye Mini Tracker, Available at <https://www.tobii-dynavox.com/devices/eye-gaze-devices/pceye-mini-access-windows-control/>, Accessed on 23/3/2018.
  - [26] Vasanth J, Jacob N, Viswanathan S. Visual function status in children with cerebral palsy, *Optometry & Visual Performance* 2014; 2(3): 251-253.
  - [27] Yang C et al., Communication Aid System for Users with Physical Impairments. *Computers and Mathematics with Applications* 2002; 43: 901-910.
  - [28] Zhai S, Morimoto C, Ihde S. Manual and Gaze Input Cascaded (MAGIC) Pointing, *ACM SIGCHI Conference on Human Factors in Computing System (CHI)*, 1999.