Voice or Gaze?
Comparative Evaluation of Vocal Joystick and Eye Tracker

Susumu Harada, Michelle Liu, Shobhit Mathur

University of Washington
Computer Science and Engineering
{harada, liujing, shobhit}@cs.washington.edu

Abstract. In this work, we present the results of our comparative evaluation of the Vocal Joystick system and the Eye Tracker system. Vocal Joystick is a non-verbal voice-based continuous mouse cursor control system that offers an alternative method of computer interaction for people with motor impairments. Our goal was to gain an understanding for how this system compares to an existing assistive technology, and we chose the Eye Tracker as our target for comparison due to its availability. The results suggest that the Vocal Joystick system can provide comparable performance as the Eye Tracker in basic target clicking tasks, and that our study participants found the Vocal Joystick to be easier to control and more preferable than the Eye Tracker system.

1 Introduction

As ubiquitous as the mouse and the keyboard has been for human computer interaction, such input devices pose a great challenge, if not an impossibility, for users with various forms of motor impairments. To enable such users to interact with computers, various forms of assistive technologies have been developed, such as the head mouse, eye trackers, and speech controlled systems [1].

As a potential alternative for hands-free computer input modality, the Vocal Joystick system was developed at the University of Washington to enable the user to continuously control the mouse cursor using their voice. The system offers promising advantages over other input methods, such as continuous control, low-cost infrastructure and minimal usage condition constraints, but the system has not yet been evaluated against other comparable technologies.

In this project, we sought to gain an understanding of how the Vocal Joystick system compares in simple target acquisition task performance and user perception of its ease of use. We were fortunate to have access to the Lab for Usability Testing and Evaluation in the Technical Communications Department, which housed an eye tracking system as well as various recording equipments. There have been other voice-based mouse cursor control techniques that have been proposed by various researchers, and we were able to obtain a copy of one of the implementations, but due to timing issues and the level of customization required for adapting it to each user, and also due to our interest in working with an eye tracker which we previously have not had the chance to use, we decided to use the eye tracker as our target for our comparative study.

We investigated the difference in user's performance and preference between the Vocal Joystick system and the eye tracker system. We started out with the following hypothesis to guide us in the design of our study:

- H1) VJ will be slower for target acquisition tasks than ET, but its error rate (false clicks) will be lower than ET.
- H2) For general browsing of web pages, VJ will be faster than ET due to lower error rate.
- H3) It will be easier to acquire smaller targets with VJ than with ET.
- H4) Users will feel more in control of VJ than ET.
- H5) Users will feel less fatigued with VJ than with ET.

We recruited participants to perform two tasks under the two modalities (VJ = Vocal Joystick, ET = eye tracker), and recorded their task times and administered a set of questionnaires to measure their subjective responses. Our results suggest that the Vocal Joystick allowed the users to perform simple target acquisition tasks at a comparable speed as the particular eye tracker we used, and that overall the participants preferred the Vocal Joystick over the eye tracker.

2 Related work

2.1 Existing mouse alternatives

There are a number of mouse alternatives that are currently available for individuals with motor impairments. Listed below are some of the representative devices that are commercially available [1]:

- Trackballs – Users move the trackball by nudging it, not necessarily grabbing it. The ball can be operated with the chin, elbow, foot or a stick held by the mouth.
- Joystick – Joysticks require a smaller range of motion than a mouse, and they can also be operated by chin or mouth movements. How far or how fast the mouse cursor moves depends on how far the joystick is pushed.
- Head pointers – Head pointers move the cursor in response to the user's head movements. They require the user to wear an infrared (Headmaster) or micro-gyroscope (Tracer) sensor on his head or perhaps a reflective dot (Tracker, HeadMouse). The system measures signals from the sensors to determine the direction on which the user's head moves. The cursor is moved in the same direction correspondingly.
- Eye Gaze and Eye Tracking – A mounted camera-like device translates eye movements and eye stares into directing the on-screen mouse.
- Speech-based Cursor Control Systems – Speech recognition tool is used for the users to control the movement of cursors.

2.2 Existing eye/gaze-based systems

Using the eyes as a source of input in "advanced user interfaces" has long been a topic of interest in the HCI field [11, 12, 13, 14]. One of the basic goals that numerous researchers have attempted to achieve is to operate the user interface through eye gaze, with pointing (target acquisition) as the core element. However, the design and implementation of eye gaze based computer input has been faced with two types of challenges. One is eye tracking technology itself. The other challenge is the human factor issues involved in utilizing eye movement for computer input. Jacob [15] discusses many of these issues with insightful observations. In search of an interaction technique that utilizes eye movement to assist the control task but does not force the user to be overly conscious of his eye movement, Zhai et al designed a technique called MAGIC (Manual and Gaze Input Cascaded Pointing) [16]. Eye-tracking has excited people outside the research community as well. Many commercial companies have developed both the hardware and software required for it [17,18]. The scope of possible applications is extensive, ranging from medical diagnostics to intuitive and fast computer interaction. Examples include:

- Computer interaction in professional environments
- Clinical diagnostics
- Security and medical imaging
- Vehicle security and vehicle interaction
- Computer gaming
- Computer interaction in a consumer environment

Today, eye-tracking already provides great value in commercial and research-related applications such as:

- Psychology and vision research
- Commercial usability and advertising studies
- Eye-based communication for people with highly limited mobility

2.3 Existing speech-based systems

Speech-based systems have recently begun to emerge both in commercial products and academic research projects. Examples of such systems that enable some form of mouse cursor control include IBM's ViaVoice [3], Dragon's NaturallySpeaking [4] and Office Organix's QPointer VoiceMouse [5].

As for academic research projects, Vocal Joystick [6] is one such system developed at the University of Washington. Vocal Joystick enables individuals with motor impairments to make use of vocal parameters to control the mouse cursor. The system uses the pronunciation of four vowels to control the movement of the mouse cursors in different directions (up, down, left and right). Vocal characteristics, such as vowel quality, pitch, and loudness are processed and extracted from individual users in real time. The system is adapted to different users in order to increase accuracy. Pitch and loudness also control the speed of cursor movement.

Other speech-based systems include SUITEKeys [7,8], which enables users to select the desired direction of movement of the cursor. The cursor then moves in that direction at a certain rate until the user issues a stop command. However, for such a system, there is a delay between the time when the user issues the stop command and the time when the system recognizes the command. To overcome this, Karimullah and Sears positioned a 'ghost' cursor to indicate where the cursor will stop after the stop command is issued [2]. The user specifies the direction and the ghost cursor moves together with the actual cursor. The user can stop the actual cursor at the desired position by issuing the stop command when the ghost cursor has reached the desired position. Moreover, systems that require the user to specify cell numbers in a grid recursively to position a cursor at the desired position have also been proposed [9]. This solution has the disadvantage of being inherently discrete. Migratory cursor proposed by Mihara et al combines discrete specifications that enable a user to specify an approximate location quickly with continuous specification that enable the user to specify a point accurately, but slowly [10].

2.4 Evaluation of speech-based systems

A thorough evaluation of speech-based cursor control systems has been conducted by Karimullah et al in [2]. They conducted both quantitative and qualitative experiments on two speech-based systems and compared them. Target acquisition tasks were designed. In the experiment, several factors that could affect the efficacy of continuous direction-based navigation, such as the distance, size, and angle of targets were investigated. Statistical data such as mean error rates, mean task completion times and standard deviations were collected. Besides quantitative analysis, questionnaires were used to perform qualitative analysis and explore users' satisfaction. In Vocal Joystick [6], the authors also conducted quantitative and qualitative studies. However, they only analyzed task completion time and number of errors. Deeper statistical analysis has not been done for Vocal Joystick system.

3 System description

3.1 Vocal Joystick

The Vocal Joystick is operated by wearing a head-worn microphone and making any one of four vowel sounds to move the cursor in the corresponding direction (Figure 1). The mapping of the vowel sounds to the movement direction is currently fixed as shown in the figure. The mouse cursor starts moving as soon as the user begins to make the vowel sound, and stops as soon as the user stops making the vowel sound. The speed of the movement is controlled by the volume of the voice, with louder sound causing the mouse to move faster. Although the version of the Vocal Joystick we used recognizes only four vowel sounds mapping to the four orthogonal directions, it is possible to move diagonally by rapidly alternating between two adjacent direction vowels (e.g. "iuiuiuiu"). Clicking is performed by making a short "ch" sound.

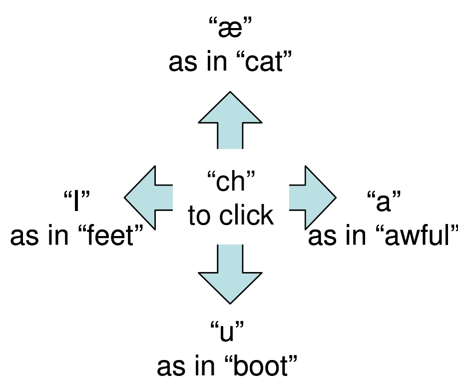

Figure 1. Vowel sounds to make in order to move the cursor in corresponding direction using Vocal Joystick

Before starting to use the Vocal Joystick, each user goes through a short adaptation phase where the system listens to the user make each of the four vowel sounds for two seconds each, in order to adjust to the specific qualities of that user's voice. The volume recorded during adaptation is also used as the reference point for the normal mouse speed.

3.2 Eye tracker

The eye tracker we used consisted of a single infrared camera that was designed to focus on the user's dominant eye, and tracks the movement of the iris by analyzing the reflection of an infrared beam emanating from the camera (information on the model of the eye tracker is provided in the experimental design section below). This particular system required the user to keep their head fairly steady, so we utilized a chin rest for the participants to rest their chin and reduce the amount of head movement. The calibration phase consists of having the user look at the center of the camera, followed by multiple reference points displayed successively on the screen. Once calibration is complete and the eye mouse control enabled, the mouse cursor follows the user's gaze. The eye tracker had an accuracy rating of 0.5 degrees visual angle, or approximately 0.5 to 1 centimeter on a computer monitor at a normal viewing distance.

Clicking is performed by dwelling, or staring at the desired point for a fixed amount of time. The dwelling time threshold is configurable, but we used the default setting of 0.25 seconds and did not modify it throughout the experiment. The actual clicking on a target was performed in two steps. The first click will invoke a zooming mode, where a fixed size window around the first click point is enlarged to occupy most of the screen, as shown in Figure 2. The bottom portion of this zoom window is a large button labeled close, which the user can click in order to dismiss the zoom window if the desired target is not within the window. Otherwise, the user clicks on the enlarged target within the zoom window to complete the action. As a feedback mechanism for indicating when the user's gaze is about to initiate a click action, the system animates a shrinking square frame around the user's gaze point when it detects that the user is fixating on a point.

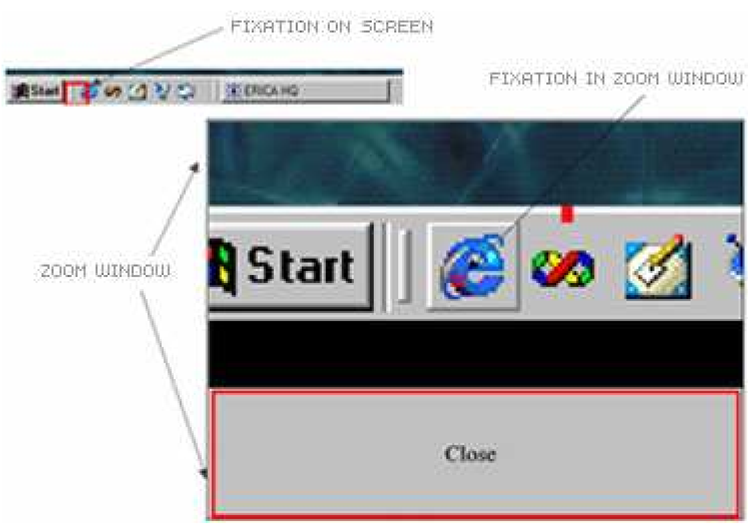

Figure 2. Zooming feature of eye tracker when clicking

4 Experimental setup

4.1 Participants

We recruited 12 participants from the Computer Science department to participate in our experiment. Of the 12, there were five females and seven males, ranging in age from 21 to 27. 7 of the participants were native English speakers and the rest of them were from Europe and Asia. 5 participants wore glasses, 2 wore contact lenses, and the rest had uncorrected vision.

4.2 Experimental design

The experiment followed a within-subject design. We exposed each participant to two different modalities: the Vocal Joystick (VJ) and the eye tracker (ET). For each modality, we had the participants perform two tasks: Target Acquisition task (TA) and the Web Browsing task (WB). The order of the tasks within each modality was fixed (TA then WB) since the two tasks are independent, but the order of the modality was counterbalanced across participants. The participants completed both tasks under one modality before moving on to the other modality.

Before starting on any task for each modality, the participants were given a description of the system they were about to use, followed by a calibration phase (for VJ, the participants were asked to vocalize the four vowels for two seconds each; for ET, the participants were asked to look at a sequence of points on the screen based on the Eye Tracker calibration software). They were then given 90 seconds to try out the system on their own to get familiar with the controls.

The experimental conditions for both modalities were shown on a 19-inch 1024x768 24-bit color LCD display. The VJ system was running on a Dell Inspiron 9100 laptop with a 3.2 GHz Intel Pentium IV processor running the Fedora Core 2 operating system. A head-mounted Amanda NC-61 microphone was used as the audio input device. The ET system was running on a HP xw4000 desktop with a 2.4 GHz Intel Pentium IV processor running Windows XP Service Pack 2. The ET camera was Eye Response Technologies' WAT-902HS model, and the software was Eye Response Technologies' ERICA version 03.08.18.

For the TA tasks, we used a Java application which we wrote that sequentially displays the starting point and the target for each trial within a maximized window and tracks the user's clicks and mouse movements. A Firefox browser was used for the WB tasks, with the browser screen maximized such that the only portion of the screen which was not displaying the contents of the web page was the top navigation toolbar which was 30 pixels high.

4.2.1 Target acquisition (TA) task

The TA task consisted of sixteen different experimental conditions with one trial each. A trial consisted of starting at a fixed region at the center of the screen (a 30 pixel wide square) and attempting to click on a circular target which appears with a randomized size and distance and angle from the center region. The sixteen conditions resulted from a fractional factorial design, based on three independent variables

1) Distance to target (100 pixels, 300 pixels),

2) Size of target diameter (20 pixels, 50 pixels)
3) Angle of approach to target (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°)

The permutation of the sixteen conditions and the corresponding target layout are shown in Figure 3. The conditions were presented to each user in a randomized order. After a target for a particular condition was successfully clicked, the start region was shown again, and once that was clicked successfully, the next target appeared. Timing data was collected for each trial between the time the start region was clicked and the time the target was clicked. We also recorded the number of times the participant clicked outside of the target.

We chose to perform a fractional factorial design instead of a full factorial design after having observed from our pilot participant that a full factorial design consisting of 32 conditions per modality took too long to perform within an hour, and our desire was to keep each experimental session within 45 minutes for scheduling purposes.

4.2.2 Web browsing (WB) task

The WB task consisted of one trial in which the user is shown a sequence of links on a web page that they need to click through and are told to follow the same links using a particular modality. The participants were first guided through the sequence by the experimenter, and then asked to go through the links themselves (by verbally recounting the links to the experimenter who then clicked the links for them) to ensure that they were familiar with the order and the location of each link. They were also instructed that if they click on a wrong link, they need to click on the browser's back button on their own to return to the previous page and try again. Once the participant was familiar with the link sequence, they were asked to navigate through those links using a particular modality. The time between when the participant started using the modality and when the participant successfully clicked on the last link was recorded, as well as the number of times they clicked on a wrong link.

The sequence of links consisted of clicking on six links starting from the CNN homepage (www.cnn.com). Most of the links were 15 pixels high (except one which was 50 pixels high) and link widths ranging from 30 to 100 pixels wide, and distances between links ranging from 45 pixels to 400 pixels, covering directions corresponding roughly to six different approach angles.

4.3 Questionnaire

After each modality, we presented the participants with a questionnaire asking them to rate the modality on a Likert scale. The questions were categorized under the following areas:

- Ease of learning how to use the system
- Enjoyment
- Fatigue
- Controllability
- Predictability (system did what the participant intended it to do)

After the end of both tasks, we also had the participants provide comparative comments on the two modalities, and solicited additional feedback through informal interviews.

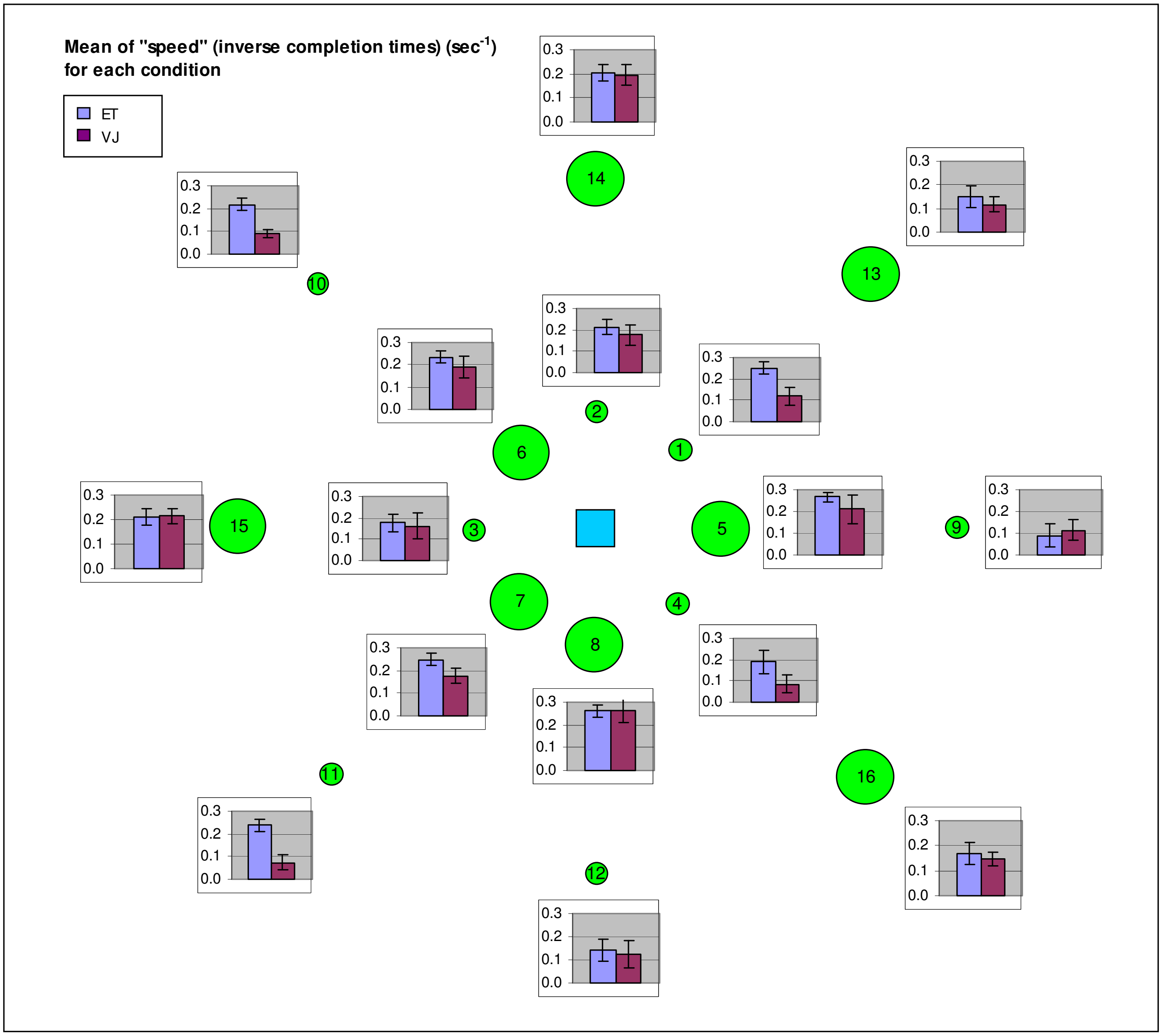

Figure 3. Mean task completion "speed" ($second^{-1}$) for the TA task across modalities.

5 Results

5.1 Quantitative tasks

The mean task completion "speed" across all participants for each of the 16 conditions across the two modalities is shown in Figure 3. We chose the inverse of the task times instead of the task times themselves as the suitable measure for examination based on residual analysis. The longer bars on the graph indicates faster performance. The circles represent the target size and distance relative to the start position (middle square) for all conditions, with the numeric labels representing the condition number. The error bars represent the 95^{th} percentile confidence interval (as with all other error bars in the other figures).

To compute the effect of various conditions on the target acquisition time we first used the correlation coefficient metric. The correlation coefficient, like the covariance, is a measure of the extent to which two measurement variables vary together. Unlike the covariance, the correlation coefficient is scaled so that its value is independent of the units in which the two measurement variables are expressed (for example, if the two measurement variables are weight and height, the value of the correlation coefficient is unchanged if weight is converted from pounds to kilograms). The value of any correlation coefficient must be between -1 and +1 inclusive. We used the correlation analysis tool in Excel to examine the effects of various variables on the target acquisition time; that is, whether large values of one variable tend to be associated with large values of the other (positive correlation), whether small values of one variable tend to be associated with large values of the other (negative correlation), or whether values of both variables tend to be unrelated (correlation near zero). We analyzed the effect of three different variables: (1) Distance: the distance to the target (Near or Far) (2) Size: size of the target (Small or Big) (3) Diagonal: The position of the target (Diagonal or along the axis). The results are shown in Figure 4.

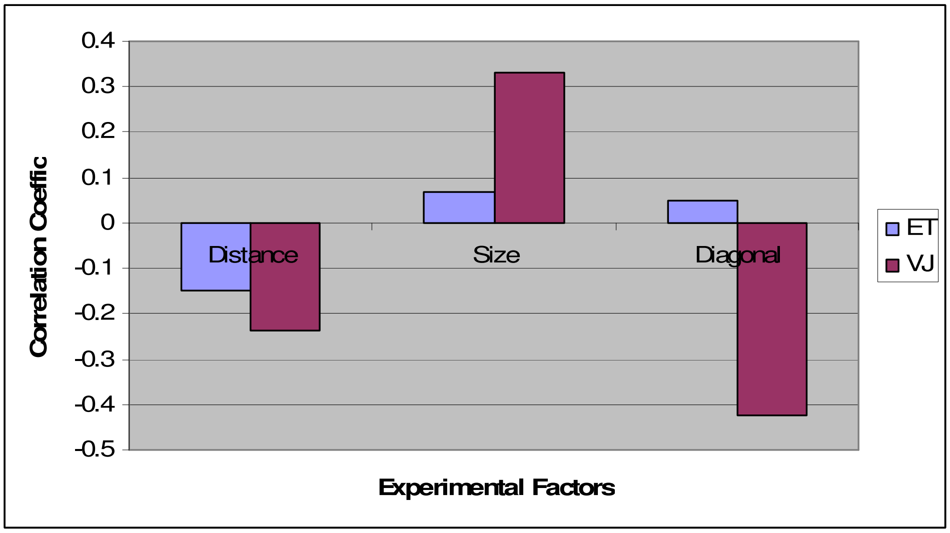

Figure 4. Correlation between target acquisition times and experimental conditions

The mean task completion times for the web browsing task are shown in Figure 5. The mean time includes any time the participant spent recovering from any errors.

5.2 Questionnaire

Figure 6 shows the means and standard deviations for the questionnaire results on completions of the target acquisition tasks and web browsing tasks.

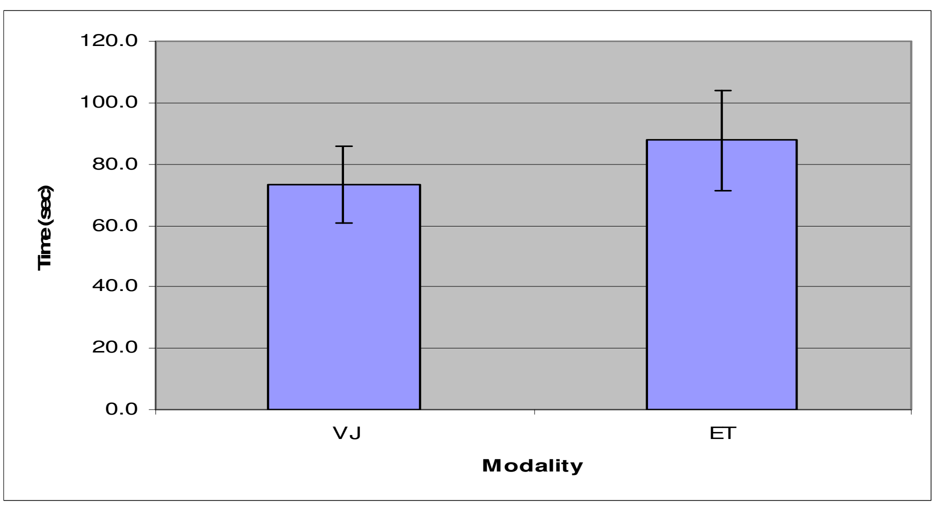

Figure 5. Web browsing task completion times (sec) across modalities

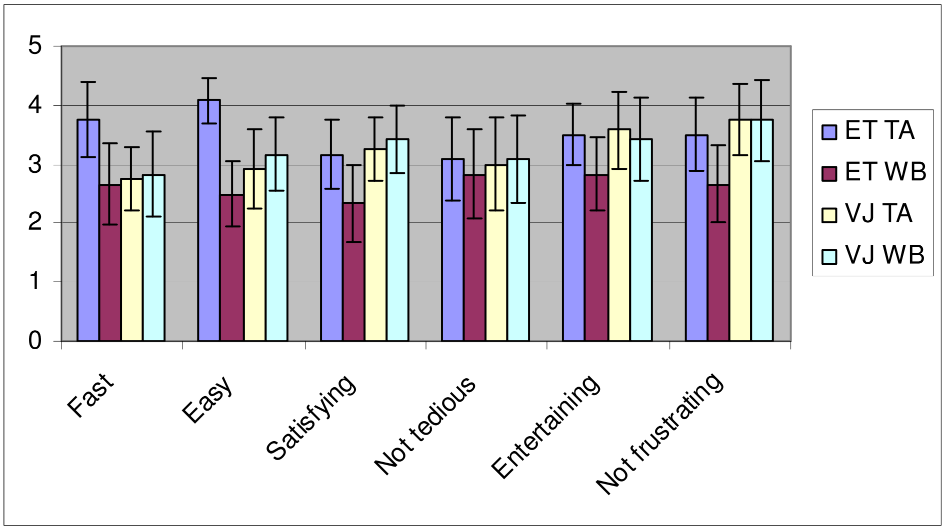

Figure 6. Mean subjective ratings for the target acquisition and web browsing tasks across different modalities. Ratings were provided on a scale from 1 to 5.

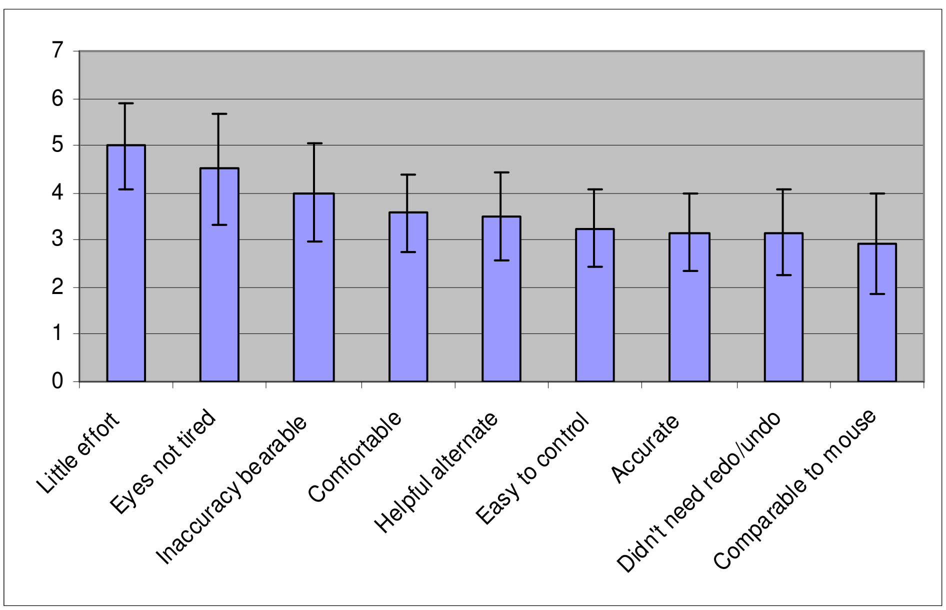

Figure 7. Mean subjective ratings on the overall performance of the Eye Tracker system. Ratings were provided on a scale from 1 to 7.

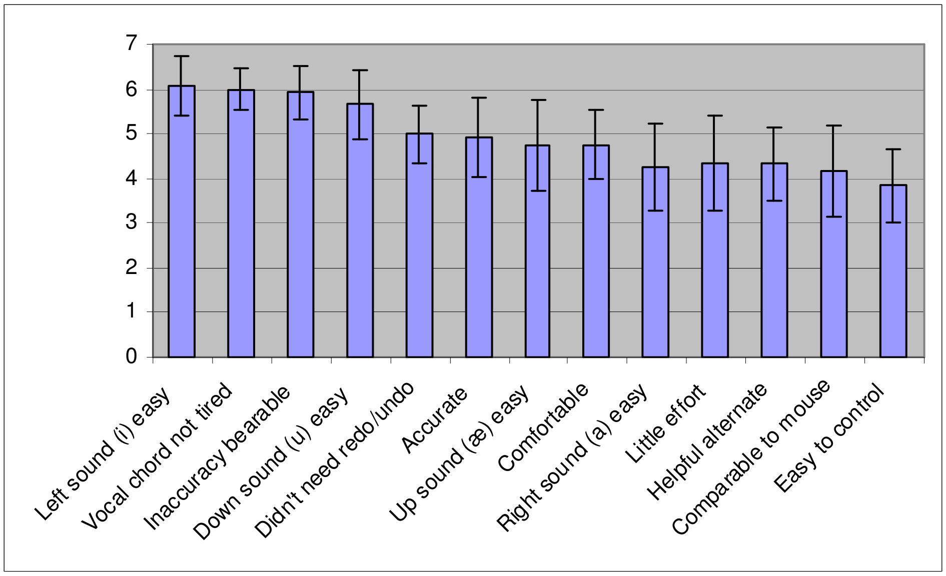

Figure 8. Mean subjective ratings on the overall performance of the Vocal Joystick system. Ratings were provided on a scale from 1 to 7.

Means responses for the questionnaire results on the overall system performances of eye tracker and Vocal Joystick are reported in Figure 7 and Figure 8, respectively.

6 Discussion

6.1 Quantitative tasks

Despite our first hypothesis, the effect of error rate was not significant with respect to the trial time (the mean number of false clicks was 1.5 for both modalities, with ET having five participants make false clicks and VJ having two).

As we can see in Figure 4, the effect of Size and Diagonal on the target acquisition speed is the most noticeable. Size is positively correlated with TA speed while Diagonal is negatively correlated. This means using VJ, bigger targets are acquired faster while targets on the diagonal take longer to be acquired. The effect of size on TA speed is expected since bigger targets are easier to acquire. VJ only moves the cursor horizontally and vertically. Hence, targets located on the diagonals are acquired slower than the ones along the axis. It is also interesting to note that Size or Diagonal does not affect the Eye Tracker TA speeds. This is because ET does not restrict the direction in which the cursor can be moved and speed does not depend on the distance of the target. We also note that as expected Distance is negatively correlated with TA speed. This effect is more pronounced for VJ than ET.

In order to see how the interaction of two factors influences the task completion time, we used repeated measures ANOVA to interpret our experimental results. We established the linear mixed effect model as follows:

$1/time \sim (C_1 \times \text{Distance} + C_2 \times \text{Size} + C_3 \times \text{Rot45} + C_4 \times \text{Rot180} + C_5 \times \text{Rot90} + C_6 \times \text{Distance:Size} + C_7 \times \text{Distance:Rot45} + C_8 \times \text{Distance:Rot180} + C_9 \times \text{Distance:Rot90} + C_{10} \times \text{Size:Rot45} + C_{11} \times \text{Size:Rot180} + C_{12} \times \text{Size:Rot90} + C_{13} \times \text{Rot45:Rot180} + C_{14} \times \text{Rot45:Rot90} + C_{15} \times \text{Rot180:Rot90}) * \text{Modality}.$

We consider the effect of various factors on the inverse of time. C_1, C_2… C_{15} represent the coefficient of corresponding variables to $1/time$ (the speed of completing the task). Variable A:B denotes the interaction of the two variables A and B. In the model, we also consider a random effect from the differences among subjects. The experiment shows that among the single variables, Distance, Size, Rot45 and Modality have a significant influence on the speed ($p < 0.001$). With regarding to the interaction of two factors, Rot45:Modality has the most significant effect on the speed ($p < 0.001$) and Size:Modality also has a significant effect ($p < 0.001$).

The biggest effects are the effects of Modality alone and the interaction of Modality with the 45 degree rotation. The results imply that in general eye tracker was faster than Vocal Joystick and the diagonal effects was more significant for the Vocal Joystick while the effect of the size difference was more important for the eye tracker.

6.2 Questionnaire

The results confirmed that participants think it is faster and easier to use Eye tracker system in target acquisition tasks while slower and more frustrating in web browsing tasks than Vocal Joystick system. In particularly, Vocal Joystick was seen to be more satisfying than the eye tracker in completing both tasks.

From these two figures, we could see that participants think Vocal Joystick requires a bit more effort to learn than Eye tracker. However, as for the easiness to control, accuracy, comfort and helpfulness, participants rated Vocal Joystick much higher than Eye tracker. From the 12 participants, 8 of them prefer Vocal Joystick system and 4 of them prefer Eye tracker system.

The questionnaire results confirmed our hypotheses. In completing the target acquisition tasks, it is faster and easier to use Eye Tracker than Vocal Joystick. While for the web browsing task, it is the converse. We think that the zooming function of Eye Tracker makes it more suitable to do target acquisition tasks. It is easier to click the targets using Eye Tracker. However, while conducting more practical tasks, such as web browsing, users not only need to click the links but also read the texts. The zooming function of Eye Tracker in this case is very frustrating. The Eye Tracker considers a long stare as a “click”, causing many false positive errors. Especially, participants complained that reading the articles was almost impossible. Moreover, when users browse a webpage, it is very natural that they will be distracted by pictures, advertisements and small articles on the page. Eye tracker in this case will be jerky. Participants liked that Vocal Joystick distinguishes between moving and clicking and this function enables them to have a feeling of being in control of the system.

As for Vocal Joystick, the pronunciation of the vowels was a bit difficult for some participants; especially participants reported that it was hard to distinguish the up sound “æ” and the right sound “a”. Moreover, for target acquisition tasks, the overshoot or the delay between the command and cursor movement causes the completion time of the tasks to be longer than that of Eye Tracker. Participants also mentioned that it is more difficult to make a diagonal movement using Vocal Joystick.

In general, participants were more satisfied with Vocal Joystick than Eye tracker. The satisfaction, less fatigue and comfort of Vocal Joystick are higher than those of Eye tracker.

7 Lessons learned

Throughout the course of designing, conducting, and analyzing the results of our experiment, we came across numerous issues which would definitely influence the way in which we will proceed in our future studies.

First, our choice of the Eye Tracker as the target for our comparative study was not investigated in great depth. We were initially attracted to it by its novelty, and as we met with several experts in the field of rehabilitation medicine and spoke with individuals with various forms of motor impairments, it became clear that the use of an eye tracker for mouse cursor control is not very common. If we had researched the use of assistive technology in greater depth at the onset of the study, we may have picked another technology such as the more common head mouse for our comparison. Also, as we discuss in our Future Work section, we could have explored other ways of rapidly prototyping and testing other voice-based cursor control systems with the Vocal Joystick. However, as most of such systems are research projects and are not widely commercialized systems, they would require more resource in setting one up for a comparative study.

We also learned the importance of starting off the design of the experiment with a good set of hypothesis and a clear indication of what measures and interactions we are interested in analyzing. We were fortunate to have access to the assistance of the consultants from the Statistics Department to help us design and analyze our quantitative and qualitative data.

8 Future work

One of the main issues in our study that we would like to address in the future is the choice of the target technology for comparison. Although we were able to gain an initial understanding of the general relationship between the Vocal Joystick and the Eye Tracker, we see that we will be able to obtain more informative and concrete comparative measures if we perform the study against one of the existing vocal mouse cursor control systems. This will allow us to reduce the number of free variables and obtain a stronger statistical validity in the quantitative analysis of task times and derive a statistical model for each system.

9 Conclusion

We have presented the motivation and the design of our comparative study of the Vocal Joystick system to the Eye Tracker system, and discussed the quantitative and qualitative results we obtained. The results indicate that the study participants were able to complete the same set of tasks with the Vocal Joystick as with the Eye Tracker, and that their performance was comparable. The questionnaire responses revealed that the participants preferred the Vocal Joystick over the Eye Tracker for ease of use and controllability. Although we recognize the limitations of our study such as the large number of free variables including Eye Tracker gaze clicking parameters and the wide range in the quality of available eye trackers, we feel that our study has provided us with an initial grounding for understanding the feasibility of the Vocal Joystick system with respect to a relatively developed input technology. We hope to extend our study by basing our choice of comparative systems on ethnographic studies and performing highly controlled experiments against other proposed voice-based cursor control systems.

10 Acknowledgement

We would like to thank all our participants for volunteering their time to participate in our study. Furthermore, we are extremely grateful to Paul Sampson, Adam Weissbarth, and Caroline Cutting of the Department of Statistics for their invaluable advice and consultation on the experimental design and data analysis.

References

1. Mouse Alternatives. http://www.techconnections.org/resources/guides/Mouse.cfm
2. A.S. Karimullah and A. Sears. Speech-Based Cursor Control. In *Proc. International ACM SIGACCESS Conference on Assistive Technologies, ASSETS '02*, pages 178-185, 2002.
3. IBM ViaVoice. http://www-306.ibm.com/software/voice/viavoice/
4. Dragon NaturallySpeaking. http://www.scansoft.com/naturallyspeaking/
5. Office Organix Qpointer VoiceMouse. http://www.officeorganix.com/
6. J.A. Bilmes, X. Li, J. Malkin, K. Kilanski, R. Wright, K. Kirchhoff, A. Subramanya, S. Harada, J.A. Landay, P. Dowden, and H. Chizeck. The Vocal Joystick: A Voice-Based Human-Computer Interface for Individuals with Motor Impairments.
7. B. Manaris, R. McCauley, and V. MacGybers. An Intelligent Interface for Keyboard and Mouse Control – Providing Full Access to PC Functionality via Speech. In *Proc. International Florida AI Research Symposium, FLAIRS-01*, pages 182-188, 2001.
8. B. Manaris and A. Harkreader. SUITEKeys: A Speech Understanding Interface for the Motor-control challenged. In *Proc. International ACM Conference on Assistive Technologies, ASSETS '98*, pages 108-115, 1998.
9. L. Dai, R. Goldman, A. Sears, and J. Lozier. Speech-Based Cursor Control: A Study of Grid-Based Solutions. In *Proc. ACM SIGACCESS Conference on Computers and Accessibility, ASSETS '04*, pages 94-101, 2004.
10. Y. Mihara, E. Shibayama, and S. Takahashi. The Migratory Cursor: Accurate Speech-Based Cursor Movement by moving Multiple Ghost Cursors using Non-Verbal Vocalizations.
11. J.L. Levine. An Eye-Controlled Computer. IBM TJ Watson Research Center: Yorktown Heights, New York, 1981.
12. R.A. Bolt. Eyes at the interface. In *Proc. 1982 Conference on Human Factors in Computing Systems*, pages 360-362, 1982.
13. C. Ware and H. Mikaelian. An evaluation of an eye tracker as a device for computer input. In *Proc. ACM SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface*, pages 183-188, 1987.

14. R.J.K. Jacob. What You Look At is What You Get: Eye Movement-Based Interaction Techniques. In *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 11-18, 1990.
15. R.J.K. Jacob. Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces. In *Advances in Human-Computer Interaction*, vol 4, Albex: Norwood, NJ, pages 151-180, 1993.
16. S. Zhai, C. Morimoto, and S. Ihde. Manual and gaze input cascaded (MAGIC) pointing. In *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 246-25, 1999.
17. Tobii Technology. http://www.tobii.se/
18. LC Technologies, Inc. Eye Gaze Systems. http://www.eyegaze.com/