A Comparative Longitudinal Study of Non-verbal Mouse Pointer

Murni Mahmud[1], Adam J. Sporka[2,3], Sri H. Kurniawan[4], Pavel Slavík[3]

[1] International Islamic University Malaysia
P.O. Box 10, 50728 Kuala Lumpur, Malaysia
murni@iiu.edu.my

[2] Hečkova 8, 14200 Praha 4, Czech Republic
adam@sporka.eu

[3] Czech Technical University in Prague
Karlovo náměstí 13, 12135 Praha 2, Czech Republic
a.j.sporka@fee.ctup.cz, slavik@fel.cvut.cz

[4] University of Manchester
PO Box 88, Manchester M60 1QD, United Kingdom
s.kurniawan@manchester.ac.uk

Abstract. A longitudinal study of two non-speech continuous cursor control systems is presented in this paper: Whistling User Interface (U^3I) and Vocal Joystick (VJ). This study combines the quantitative and qualitative methods to get a better understanding of novice users' experience over time. Three hypotheses were tested in this study. The quantitative data show that U^3I performed better in error rate and in simulating a mouse click; VJ was better on other measures. The qualitative data indicate that the participants' opinions regarding both tools improved day-by-day. U^3I was perceived as less fatiguing than VJ. U^3I approached the performance of VJ at the end of the study period, indicating that these two systems can achieve similar performances as users get more experienced in using them. This study supports two hypotheses but does not provide enough evidence to support one hypothesis.

Keywords: Voice-based interface, non-verbal vocal input, speech recognition, cursor control, continuous input, mouse cursor, acoustic gestures.

1 Introduction

Most recent days' interactive systems follow the WIMP (Windows, Icons, Menus, and Pointer) paradigm and can usually be controlled optimally using a mouse or equivalent pointing devices such as a trackball or touchpad. However, people with some impairments of upper limbs, especially those who have problems performing fine motor movement with their fingers, face difficulties when using these devices, and have to use alternative devices to accommodate their needs and capabilities. Many of these devices are based on dedicated hardware solutions, such as sip-andpuff controllers, feet-operated input devices, or eye trackers. Being typically produced in small quantities, they can be very costly and therefore are not affordable to some users.

Increasing reliability and deployability of the speech recognition on the lower-end computer systems make this technology a promising alternative. However, the speech recognition systems operate on a query—response basis, where the system usually waits for the user to complete the utterance before responding. This makes the speech recognition inconvenient to use in real-time continuous tasks, including mouse pointer movement, where the minimum delay of the system is a critical feature of the feedback loop of the system [22].

In the last decade, the non-speech (or non-verbal) input has started to emerge. It is based on production of sounds other than speech by the user's vocal tract. It can be considered a counterpart to the non-speech output modality, which is mostly used in presentation of data (such as [3, 7]). In non-speech input, the sound is analyzed and certain features are extracted, such as the pitch profile, or the sound timbre, in order to solicit information from the user. Depending on the application, these features may be assigned to different elementary controls of the user interface.

Recently, this form of input has been employed in several tools that implement a virtual mouse device. This paper aims at comparing two of these systems, Whistling User Interface (U^3I) [19] and Vocal Joystick (VJ) [10]. These two systems will be explained in more detail in Section 4.

2 Motivation

Many usability studies have pointed out that there is a considerable difference between involving so-called novice or expert users because these users may have different levels of experience with the system being evaluated. Therefore, it is important to study users over time as they develop expertise in using the systems to answer a key question of how the user's experience of a system's usability changes when they transform from being novices to being more expert, if usability problems really disappear over time when users get more familiar with a system [13]. This forms the motivation of the longitudinal nature of the reported study.

This study combines quantitative and qualitative methods to arrive at a better understanding of the usability and performance of the two tested systems by novice users, as it develops over time. This combination is expected to be able to complement the pictures provided by individual methods in regard to user performance and opinions on the tested systems.

Non-speech voice input as a user interface control modality holds the potential for offering effective input modality for individuals with physical or situation-induced motor impairment, but this space has been relatively unexplored. Through studying the adoptability of such techniques by novice users, we hope to gain a better understanding of whether such systems are indeed viable, and what their ease of use and learnability are. By combining results from prior studies that compare these systems to common input devices (e.g. VJ and the mouse in [4]), we can also place these novel input methods on the map of other existing input devices.

We decided to investigate user experience on VJ and U^3I for several reasons:

- Past studies had compared the performances and user opinions of one of the systems (VJ) with common input devices, and therefore, the comparison between VJ and U^3I can be placed on a 'bigger map' of input devices in general.
- Both systems share certain properties (based on non-speech vocal input and respond immediately to changing features of the sound), thereby minimizing the variability when investigating the causes of performance and opinion differences.

There are three hypotheses that this study tests (the reasoning behind these hypotheses will be elaborated in the sections that explain how these two tools work):

<u>H1:</u> U^3I is faster than VJ in emulating a click (reasoning: humming is easier to produce than the sound 'k' used by VJ to emulate a click).

<u>H2a:</u> User performance is higher using VJ as opposed to using U^3I (reasoning: it is easier to associate the sounds and movements produced in VJ than it is in U^3I).

<u>H2b:</u> User opinion of U^3I is better than that of VJ (reasoning: it is less tiresome to produce 'mmmm' than it is to produce 'aaah', 'iiih', 'uuuh', 'eeeh').

<u>H3:</u> User performance and opinions regarding both tools would improve at the end of the study period.

3 Related Work

Non-speech input has been evaluated in many different contexts, such as computer games [10], interactive art [1], music training [17], or keyboard emulation [20]. In this section, an overview of the voice-based methods of mouse pointer control is given.

Non-speech Methods. Only non-verbal sounds are used to control the mouse pointer.

- Whistling User Interface (U^3I) [19] is based on the use of tones (in whistling or humming) where the difference between the initial pitch and current pitch determines the speed of motion.
- Vocal Joystick (VJ) [10] is based on the assignment of different vowel sounds ('aah', 'eeh', etc.) to four or eight basic directions. The movement continues for as long as the sound is being produced. The loudness governs the speed of motion.

Hybrid Methods. The hybrid methods make use of speech commands that are augmented by non-verbal vocalizations.

- Non-verbal quantification of speech-issued commands is proposed in [11]. The users would utter 'move down' and then produce sound such as 'aaah' for as long as they wished the movement to last.
- Migratory Cursor [16] combines the speech recognition for coarse approach of cursor towards the target and subsequent refinement of the position with a non-verbal sound.

Speech-based Methods. In speech-based methods, the mouse cursor is controlled by speech utterances only.

- Direction commands-based methods—such as SUITTEKeys [15] interface—make use of discrete speech commands such as 'move mouse left' (initiating motion in that direction), 'stop' (motion is stopped). The Dragon NaturallySpeaking®

package, in addition to these basic commands, allows control the cursor speed by commands such as 'faster', 'much faster', etc.

- Grid-based systems, such as [12, 5] are based on recursive subdivision of screen by a grid, typically 3 by 3 cells, assigned numbers 1 through 9. Upon selecting a cell by uttering a number, the cell becomes subdivided a similar manner. As soon as the focus is over the target, the user may utter a 'click' command.
- Voice Mouse [8] uses a similar assignment of vowels as VJ. However, the system requires the user to produce one sound to initiate the motion and another to end it.

4 Description of the systems

The Vocal Joystick System. The Vocal Joystick (VJ) system [10] offers as one of its modes a cursor control mode that allows a user to move the mouse pointer by making vowel sounds that have been assigned to each direction. Figure 1 shows the mapping of the vowel sounds to the directions for the 4-way version of VJ, which was the version that was used in our experiment. An 8-way version is also available. A more detailed technical description of the VJ system can be found in [4].

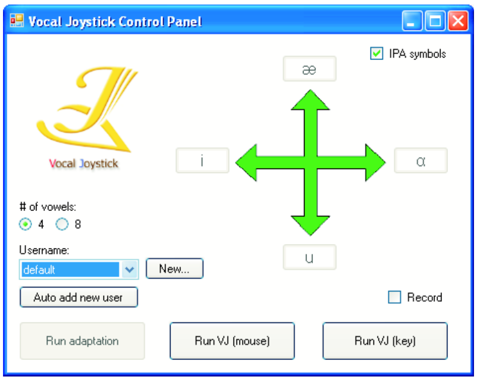

Fig. 1. Set-up Screen of the Vocal Joystick Application, showing the assignment of sounds to directions of mouse pointer movement.

To move the mouse cursor, the user starts vocalizing one of the vowel sounds, and the cursor would immediately begin moving in the corresponding direction with a speed that is proportional to loudness. The user can continuously vary the vowel or the loudness, and the cursor stops moving as soon as the vocalization terminates. Clicking is achieved by making a short discrete sound (in this case the consonant 'k'), and toggling of the mouse button for dragging is performed by making the discrete sound 'ch'. The control panel of VJ is shown on Fig.1.

VJ can be used without any user specific training, but its performance can be improved by adapting the system to each user's vocal characteristics. This process involves the user vocalizing each of the vowel sounds for two seconds at their normal loudness, and the actual adaptation step takes less than a second.

The recognition of the vowels is very robust and accurate compared to recognizing words under conventional speech recognition systems.

Whistling User Interface. This system (U^{3}I) allows the users to operate the mouse by producing a series of acoustic gestures [19]. Depending on the length and pitch profile (melody) of the gesture the system moves the cursor in specific direction at certain speed or emulates a click.

The cursor may be moved along one of the x- and y-axis at a time. Depending on the initial pitch of the tone (either below or above a specified threshold), the gesture would drive the cursor either horizontally or vertically. A mouse click is emulated when a short tone is produced by the user. All gestures are shown in Fig. 2.

Originally, only whistling was received by the system, which gave the system its name. The analysis of humming (i.e. producing ‘mmm’ sound) was implemented as a user test [21] proved that humming was easier to produce and less tiresome for the users.

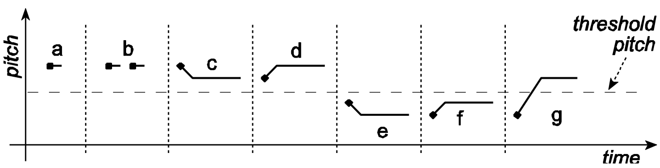

Fig. 2. Example gestures of the U^{3}I system. a – click, b – double-click, c – downwards, d – upwards, e – to the left, f – to the right, g – a movement to the right.

The system runs as a standalone application on win32 platforms (see Fig. 3). Its user interface comprises of a single dialog window which also provides a visual feedback of the system for set-up purposes (checking proper function of the microphone, calibration of parameters, etc.) However, while used in normal operation, the system provides no additional feedback than the one provided by a common mouse.

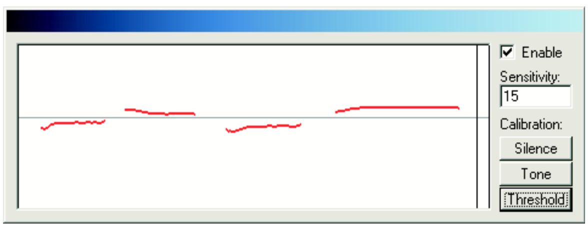

Fig. 3. User Interface of U^{3}I Application

5 Experiment

Participants. Ten participants were recruited from the university, high school and members of the public in the Manchester Area (mean age=18.8, SD=9.1). All participants (8 female, 2 male) have some computer and mouse experience (mean=3.2 years, SD=2.6). None of the participants was familiar with the target acquisition task. All participants signed informed consent forms before the experiment.

Apparatus. The experiment was conducted on a Dell Inspiron 630m laptop with 1.6 GHz Intel Centrino processor and 1 GB of RAM running the Windows XP Home operating system. The resolution of the 14.1" LCD was 1280 by 800 pixels. The experiment was conducted in full screen mode, with the user's head situated about two feet from the screen. A noise-canceling headset microphone (Altec Lansing AHS322) was used for audio input.

Design. A series of controlled experiments were executed longitudinally over five consecutive days. The experiment sessions were conducted at about the same time of day everyday. The study followed a 2×2×2×8 within-subjects design with repeated measures (3 trials per condition). The arrangement of the tools and the size of the target were counterbalanced. It was decided to run the experiment with an index of difficulty (ID in bits) of 3 (replicating one of the IDs used in a prior target acquisition work comparing VJ's and mouse's performances [4], as the results of this study will need to be extrapolated with mouse performance). The factors and levels were:

- Modality (M) {VJ, U^3I}
- Target shape (T) {bar – similar to the stimulus used in Fitts' 1954 experiment [6], circle}
- Target width-distance (to maintain ID=3, in pixels) {W=50, A=350; W=100, A=700}
- Approach angle (θ in degrees, counterclockwise) {0 = the direction to the right of the screen, 45, 90, 135, 180, 225, 270, 315}
- Day (D) {1..5}

Procedure. The experiment on day 1 consisted of the following (in the order it was conducted):

1. A demographics questionnaire to elicit age, experience with computer, experience with speech-based tools and experience with standard input devices.
2. Calibration and practice of the two modalities.
3. Training sessions with both modalities. Each participant performed 16 target acquisition trials for each tool and each of the bar and circle targets. A stimulus used in the experiment is shown in Fig. 4.
4. 48 target acquisition trials for each modality and each target shape. The sequence of modalities and target shapes were counterbalanced across participants.
5. Subjective rating of each modality in terms of their ease of learning, ease of use, level of fatigue, level of frustration, satisfaction and confidence.

On days 2-5 the participants performed activities 2, 4 and 5. In addition, on the last day the participants also performed a simple reaction time (SRT) task aimed at measuring how the two modalities process sounds intended to simulate a click.

The SRT test was performed using an online stimulus called The Online Reaction Time Test [2]. This was a simple operation where the participant clicked the right button to start, which caused the red light to turn on. The participant had to click this button as fast as they could as soon as the red light turned green. No cursor movement was measured in this experiment. Each participant was asked to click 20 times for each tested system. After 20 times, the data were analyzed to remove outliers. If there were outliers, the participants were asked to click some more times to make up 20 useful data points. The data were then averaged. To assist users, a printout of the mapping between the voices and movements was provided throughout the experiment to guide the participants (shown in Fig. 5).

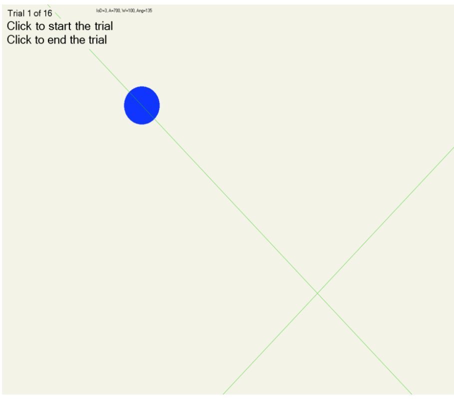

Fig. 4. A screenshot of a stimulus. The cursor position is at the crossing of the two diagonal lines.

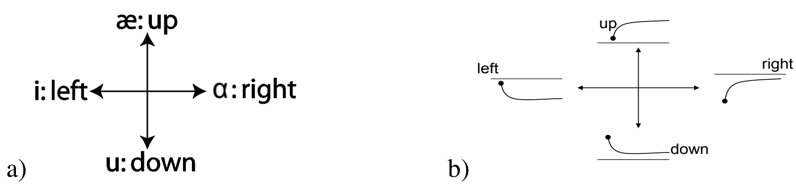

Fig. 5. Printout of mapping used in the experiment for (a) VJ and (b) U^{3}I.

6 Results

In the rest of the paper, data we report as significant are at $p \leq .05$ level.

6.1 Quantitative Results

The error rate (percentage of the trials in which the participant failed to acquire the target) over the five days for each of the modality is shown in Figure 6. Trials in which the participants failed to acquire the target were not included in subsequent analyses.

Figure 7 shows the movement times for each modality across the five days. The data had to be analysed separately for *cardinal* (horizontal or vertical) and *ordinal* (oblique) directions, as the index of difficulty of the ordinal targets were higher than those of the cardinal targets. Helmert contrasts showed that the day effect was not significant after day 4 for U3I, but that it was not significant at all throughout the five day period for VJ, in support of H2b.

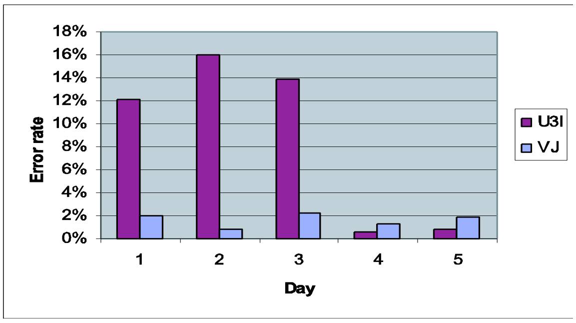

Fig. 6. Error rate for each modality over five days.

In order to characterize and compare the movement of the pointer under each modality as it traverses from the origin to the target, we analyzed six different measures of mouse pointer movement as described in [14]. Figure 8 shows the plot of the measures for the two modalities grouped by target direction categories. The data was aggregated over the five days since no learning effect was observed for any of these six measures. The data on six measure support H2a (User performance using VJ is better than those using U3I).

- Target Re-entry (TRE) is the number of times the pointer moved from within the target area to outside the target area and back in. This is a measure of overshoots, and Figure 7(a) shows that the TRE remains constant for VJ regardless of whether the movements are ordinal or cardinal, U3I's TRE was significantly higher than VJ for both movements, with the difference between the two system's performances becoming more dramatic with the ordinal movement.
- Task Axis Crossing (TAC) is the number of times the pointer crossed the line connecting the starting point to the target centre (the task axis). It is expected that TAC is smaller for cardinal movements as the users only need to maintain the movement in one direction. However, in two directional movements, TAC shows that users have more problems in 'wobbliness' when using U3I compared to when VJ was used.
- Movement Direction Change (MDC) is the number of points along the pointer path where the change in direction is parallel to the task axis.
- Orthogonal Direction Change (ODC) is the number of points along the pointer path where the change in direction is perpendicular to the task axis. MDC and ODC represent similar measures, the degree of overshoots along the task axis. It is interesting to note that users seem to have more problems with MDC than with ODC when using VJ at ordinal movement while the performance was pretty much constant when using U3I. This might indicate that certain movement change using VJ is more difficult to make than other change (e.g., changing from 'aah' to 'uuh' might be more difficult than changing from 'uuh' to 'iih').
- Movement Variability (MV, in pixel) is the standard deviation of the distances of the points along the path to the task axis from their mean.

• Movement Error (ME, in pixel) is the average distance of the points along the path from the task axis. Both MV and ME are a detailed measure of wobbliness, and therefore these two diagrams depict similar trends to those of TAC. VJ was significantly lower than U³I on all measures regardless of the target direction.

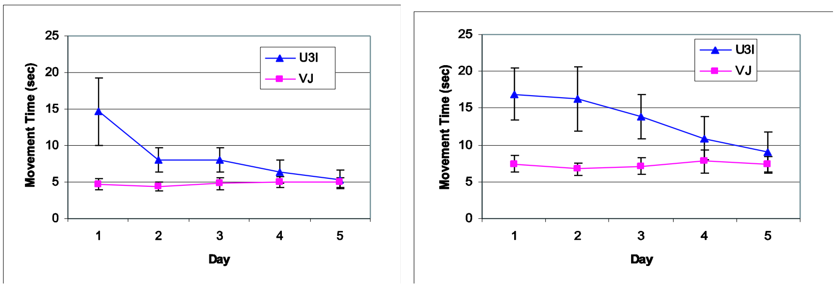

Fig. 7. Movement times over five days for VJ and U³I for targets along (a) cardinal directions and (b) ordinal directions. The bars show the standard deviations.

6.2 Simple Reaction Time

This experiment aimed at testing H1 (U³I is faster than VJ in emulating a click as humming is easier to produce than 'k' used by VJ to emulate a click). The results suggest that the U³I was significantly faster in emulating a click (mean=578ms, SD=84) than VJ was (mean=857ms, SD=129) with ($F(1, 18)=32.744$, $p < 0.01$). Therefore, the empirical data gathered in this study is in support of H1.

6.3 Subjective Ratings

Both tools were rated at the end of each day session on six aspects: Q1 through Q6 (Q1 = ease of learning, Q2 = ease of use, Q3 = level of fatigue, Q4 = level of frustration, Q5 = satisfaction and Q6 = confidence), with 1 = lowest and 5 = highest. For analysis, the ratings for fatigue and frustration were reversed so that lower numbers indicate better opinions regarding fatigue and frustration. Figure 9 and 10 shows the progress of ratings over five days. Both figures show that users' opinions regarding both systems improved by days, in support of H2b.

The changes in ratings by day suggest that overall the participants' perception of the two tools had improved. For each aspect, the results suggest that VJ received significantly better rating than U³I as tested through t-tests. This suggests that there is not enough evidence to support H3.

However, the Table 1 shows that the improvement in user perception was not equal for both tools. The t-test analysis comparing day 1's and day 5's perceptions show that the mean ratings for both tools were not significantly different for ease of learning (Q1). The change in ease of use (Q2) was not significant for U³I. The change in frustration (Q4) was not significant for VJ; with VJ receiving a higher rating than U³I did for this aspect.

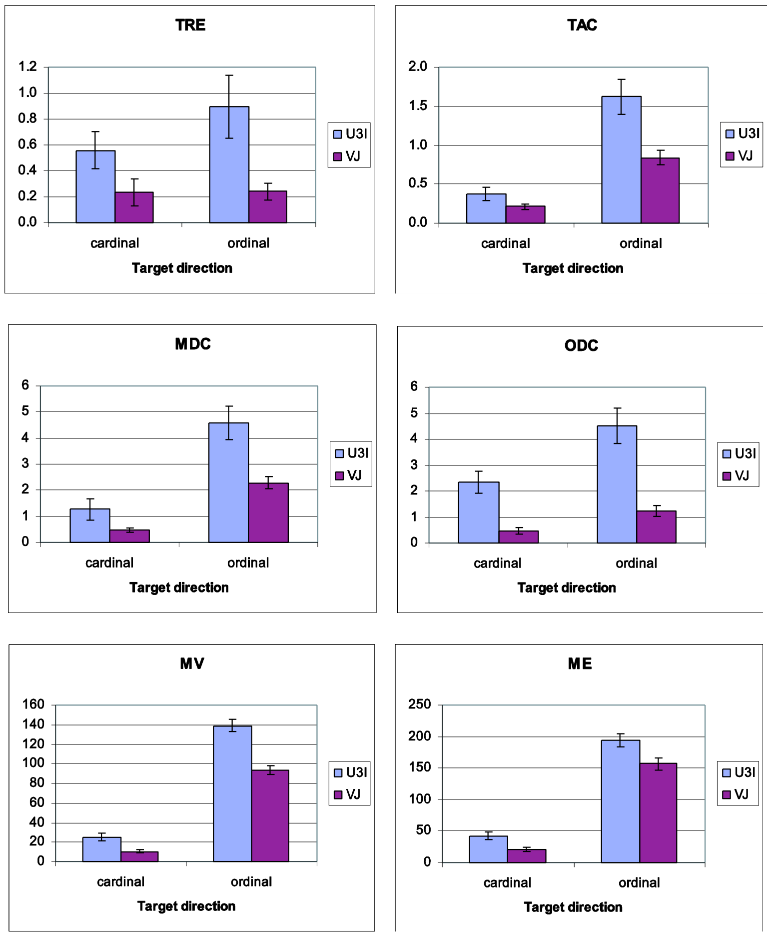

Fig. 8. Mouse pointer trail movements for VJ and U^{3}I grouped by directions.

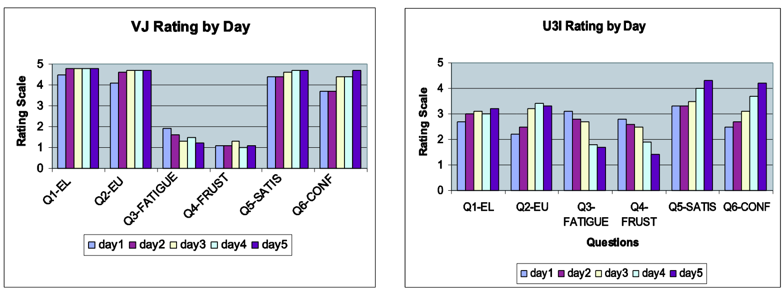

Fig. 9. Subjective rating across five days for (a) VJ and (b) U^{3}I.

Table 1. t-test result of comparing day 1 and day 5 for each aspect (df=9).

	VJ		U^3I	
Aspect	T	sig	t	sig
Q1	-1.406	0.193	-1.103	0.299
Q2	-3.674	0.005	-2.181	0.057
Q3	2.689	0.025	4.583	0.001
Q4	same rating		2.941	0.016
Q5	-1.406	0.193	-2.372	0.042
Q6	-2.372	0.042	-5.075	0.001

6.4 Qualitative Data

The following qualitative data were derived from a combination of the participants' comments and the researchers' observation. Participants were generally optimistic about both systems.

Production of sounds. When using VJ, some of the participants had problems with articulating the appropriate vowel sounds. In situations where the participants were not able to replicate the sounds stored in the existing VJ profile (causing the cursor to move in undesired direction), a new profile was created.

The humming sounds in U^3I were much less challenging for the users in terms of articulation but people with pitch hearing would be in advantage when using U^3I. There key problem was for the users to memorize the threshold pitch. Users had to reselect their threshold pitch everyday, making U^3I sessions, in general, to last longer than VJ sessions.

Fatigue. The participants generally reported the VJ to be more fatiguing than the U^3I due to the need of keeping the proper articulation of the vowel sounds, especially when moving the cursor over large distances.

Cursor control. When using VJ, it was observed that when the cursor was close to the target, the users tended to decrease the volume to avoid overshooting. It was also observed that some participants could not reach the target in one continuous movement because of short of breath. One participant commented that there was a noticeable delay between the voice and the corresponding cursor movement.

In situations where the participants overshot the target and the cursor went off the screen, sometimes they could not rectify the problem even after producing various voices in an attempt to bring the cursor back. In these situations, the experimenter would use the mouse to reveal the position of the cursor.

One suggestion that the participants made was to implement a filter of unwanted noises. These include background noise (to which both systems were very sensitive) and vocalizations made by user other than the prescribed ones.

Ease of learning and use. Many participants mentioned that it was difficult to use the systems at the beginning but with practice, they were able to gain their skills. Participants were generally in agreement that VJ was easier to learn than U^3I. It was found more comfortable for voicing out vowels was a familiar activity for the users. Though humming sounds in U^3I were easier to produce, memorizing the threshold pitch was a hindrance. The voice—movement mapping printouts were helpful by most participants. The participants suggested improving both tools by providing a more user-friendly interface and cues that would help them recover from errors.

7 Discussion

The analysis of the SRT experiment supports H1, the users were faster when using U^3I. This suggests that producing a tonal gesture (i.e. 'mmm') is more practical and easier to pronounce than the consonant 'k'. Users were more frustrated when they were unable to produce the correct sound. As VJ is affected by the volume of the voice produced, when the users raised their voices, the sound 'k' was easily distorted and thus not recognized by VJ. The users reported that they felt that more extensive movement of lips was required to click using VJ than using U^3I.

Comparing the learning curves for the cardinal versus ordinal directions, it is interesting to note that the curve for U^3I is more gradual for the ordinal directions. This seems to reflect the observation that given the 4-way nature of the two input methods used, moving in a diagonal can take more getting used to than simply moving along one of the 4 directions, as the user will need to constantly change directions using the appropriate sounds.

Another point of intrigue is the fact that for the VJ, almost no learning effect was observed. This might suggest that the participants were able to get used to the VJ control very quickly and therefore plateaued, or that the learning curve is very gradual and that with longer usage time, their performance will improve further.

Comparing the two modalities, it appears that U^3I approached the performance of VJ at the end of the five day period, and suggests a need for a prolonged study to assess whether one would significantly outperform the other.

From the experiment, it can be concluded that:

- Overall, the participants' opinions improved by day for both systems (supporting H3 and highlighting the importance of longitudinal study).
- U^3I is significantly better than VJ in emulating mouse clicks (supporting H1).
- VJ was better than U^3I in various performance measures (supporting H2a).
- VJ received better user opinions than U^3I in various performance measures (indicating that there is not enough evidence to support H2b).

The results show that the participants' objective measures and subjective opinions improved longitudinally. One interesting thing to note is that there was contradiction between the ratings and comments for fatigue. Some participants stated that VJ was more tiring than U^3I as it required more lip movements. However, the quantitative ratings suggested that VJ was less tiring than U^3I.

One possible explanation was that, whilst VJ did require more extensive lip movement, there were more overshoots when controlling the cursor with U^3I, causing the users to perceive U^3I as more tiring in their ratings.

8 Conclusion

This paper presents a longitudinal study of two cursor control systems that share some characteristics: Both systems are non-speech and continuous. The observation of novice users over time through a mixed of qualitative and quantitative methods shows the change in user experience as they gain more expertise in using the systems.

The study reveals that the participants' opinions on both tools improved day-by-day. Both U^3I and VJ excelled in some aspects, but the performances of both systems became rather similar toward the end of the study period. Combining this finding with the finding from a previous study comparing VJ and speech-based cursor control system and a hand-operated joystick [10], where VJ can potentially approach the performance of a hand-operated joystick with expert users, we can argue that these two systems have the potential for being used as a standard input device (although as [10] shows, the performance is still three times slower than when a mouse is used).

There are some limitations of the study. In order to compare the user performance with standard input devices, this study should be interpolated with another study [10]. In future studies, it is important to ensure that mouse performance is captured. Secondly, only 10 users were involved in the study, over a period of 5 days (even though each user performed 96 trials per day).

A longer longitudinal study with more users would reveal more interesting findings, such as stronger evidence of performance plateauing, especially with U^3I. Finally, as this system would benefit user population with upper limb impairment the most, future studies should include this population to investigate their opinions and acceptance regarding these two systems.

9 Acknowledgements

We would like to thank Susumu Harada of University of Washington for providing the Vocal Joystick system, the platform for performing the target acquisition tasks, and his valuable advice on data analysis. We would also like to thank all the participants who volunteered to take part in the user evaluation.

This work has been partly supported by the Ministry of Education, Youth and Sports of the Czech Republic under the research program LC-06008 (Center for Computer Graphics).

References

1. Al-Hashimi, S.:. Blowtter: A voice-controlled plotter. In: Proc HCI 2006 Engage, The 20th BCS HCI Group conference in co-operation with ACM, vol. 2, London, England, 2006.
2. Allen, J.: The online reaction time test. On-line http://getyourwebsitehere.com/jswb/rttest01.html, retrieved August 2006, 2002.
3. Barra, M., Cillo, T., Santis, A. D., Petrillo, U. F., Negro, A., Scarano, V., Matlock T. Maglio, P. P.: Personal webmelody: Customized sonification of web servers. In: Proc 2001 International Conference on Auditory Display, August 2001.

4. Bilmes, J. A., Li, X. , Malkin, J. , Kilanski, K. , Wright, R. , Kirchhoff, K., Subramanya A., Harada S., Landay J. A., Dowden P., Chizeck H.: The Vocal Joystick: A voice-based human-computer interface for individuals with motor impairments. In: Proc Human Lang Tech Conf./Conf. on Empirical Methods in Natural Language Processing, October 2005.
5. Dai L., Goldman R., Sears A., Lozier J. Speech-based cursor control: a study of grid-based solutions. SIGACCESS Access. Comput., (77-78):94–101, 2004.
6. Fitts, P. M.: The information capacity of the human motor system in controlling the amplitude of movement. J of Experimental Psychology, vol 47, no 6, 1954, pp 381–391.
7. Franklin, K., Roberts, J. C.: Pie Chart Sonification. In Proceedings Information Visualization, 2003, pp. 4–9. IEEE Computer Society
8. de Mauro, C., Gori M., Maggini M., Martinelli, E.: Easy access to graphical interfaces by voice mouse. Technical report, Universita di Siena, 2001.
9. Hämäläinen, P., Mäki-Patola, T., Pulkki V., Airas M.: Musical computer games played by singing. In Proc of 7th Intl Conf on Digital Audio Effects, Naples, pp 367–371, 2004.
10. Harada S., Landay, J. A., Malkin, J., Li, X., Bilmes, J. A. The vocal joystick: evaluation of voice-based cursor control techniques. In: Proc ASSETS 2006, Portland, Oregon, 2006.
11. Igarashi, T., Hughes J. F.: Voice as sound: using non-verbal voice input for interactive control. In UIST ’01: Proc 14th Annual ACM Symp on User Interface Software and Technology, pages 155–156, New York, NY, USA, 2001. ACM Press.
12. Kamel, H. M., Landay J. A.: Sketching images eyes-free: a grid-based dynamic drawing tool for the blind. In: Proc ASSETS 2002, New York, pp 33–40, 2002.
13. Kjeldskov J., Skov M. B., Stage J.: Does time heal?: a longitudinal study of usability. In OZCHI ’05: Proceedings of the 19th conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction, pages 1–10. Computer-Human Interaction Special Interest Group (CHISIG) of Australia, 2005.
14. MacKenzie, I. S., Kauppinen, T., Silfverberg, M.: Accuracy measures for evaluating computer pointing devices. In Proc CHI ’01, pp 9–16, New York, 2001. ACM Press.
15. Manaris, B., McCauley R., MacGyvers V: An intelligent interface for keyboard and mouse control – providing full access to PC functionality via speech. In Proc 14th Intl. Florida Artificial Intelligence Research Society Conference, pages 182–188. AAAI Press, 2001.
16. Mihara, Y., Shibayama E., Takahashi S.: The migratory cursor: accurate speech-based cursor movement by moving multiple ghost cursors using non-verbal vocalizations. In Proc ASSETS 2005, pp 76–83, New York. ACM Press.
17. Nakano, T., Goto, M., Ogata, J., Hiraga, Y.: Voice drummer: A music notation interface of drum sounds using voice percussion input. In Proc. of UIST 2005, pp 49–50. ACM Press.
19. Sporka A. J., Kurniawan, S. H., Slavík, P.: Accoustic control of mouse pointer. Universal Access in the Information Society, 4(3):237–245, 2006.
20. Sporka A. J., Kurniawan, S. H., Slavík, P.: Non-speech operated emulation of keyboard. In J. Clarkson, P. Langdon, and P. Robinson, editors, Cambridge Workshop on Universal Access and Assistive Technology, CWUAAT 2006. Designing Accessible Technology, pages 145–154. Springer-Verlag London Ltd, 2006.
21. Sporka A. J., Kurniawan, S. H., Mahmud, M., Slavík, P.: Tonal control of mouse cursor (U^3I): A usability study with the elderly. In Proc HCI International 2005. Lawrence Erlbaum Associates, July 2005.
22. Sporka A. J., Kurniawan, S. H., Mahmud, M., Slavík, P.: Non-speech input and speech recognition for real-time control of computer games. In The Proceedings of The Eighth International ACM SIGACCESS Conference on Computers & Accessibility, ASSETS 2006, Portland, Oregon. ACM, 2006.