Does Eye Gaze combined with EMG switching outperform traditional Eye Gaze AAC systems?

Lester Lo
Biomedical Engineering Student
RMIT University
Melbourne, Victoria, Australia

Abstract

This study was designed to test the speed and accuracy of an eye gaze augmentative alternative communication (AAC) system using 'dwell' selection (extended gaze), compared to eye gaze combined with electromyography (EMG) switching for selection. The results presented a 47% increase in typing speed when using eye gaze with EMG switching as opposed to eye gaze with dwell. Subjectively, the participants found that the eye gaze with EMG switching was less fatiguing, less stressful and induced less mental overload than eye gaze with dwell. In addition, the participants found that they could operate the eye gaze with EMG switching system better and more naturally. The accuracy rates for the eye gaze with dwell and eye gaze with EMG switching were not significantly different at 99.05% and 98.83%, respectively. Overall, study participants preferred the use of eye gaze with EMG switching over traditional eye gaze with dwell.

1. Introduction

Eye gaze technology is a method of accessing augmentative and alternative communication (AAC) systems suited for users with specific motor diseases or physical disabilities that restrict their independence and ability to communicate with others. The system works via a camera that tracks the path of the user's gaze, connected to and controlling a device as seen in *Figure 1*, much as a mouse would do with a computer. In order to select the item being looked at, the user can either use an extended gaze (dwell) or a switch to signify the intended use of the button, as opposed to general viewing.

Figure 1: Render of a model using an eye gaze system

When using extended gaze to select, the time elapsed between looking at and selecting the button is called the 'dwell time'.

The dwell time can be adjusted depending on the preference and experience of the user, as there are advantages and disadvantages that come with long or short times. Dwell times that are too short may provide quick and snappy responses but also can be a chaotic and stressful experience for the user, with more errors due to unintentional activations [1, 2]. Excessively long dwell times on the other hand, may work to reduce these unintentional activations, but extended staring on each item can induce significant eye fatigue for the user [3]. The dwell time is typically around 500-1000ms (0.5-1 second), with some users able to reduce dwell time to 200-300ms without compromising accuracy [4]. As a result, these methods of typing can often be slow, as it requires the user to wait for their dwell time to elapse before moving on to the next letter [4]. Despite full cognitive function, a ten-character word with a 1000ms dwell time will require an excess of 10 seconds to be completed by the user as time is required to cognitively process the next key to select and its location on the keyboard [4]. Consequently, dwell time typing is generally slow, often less than 10 words per minute (wpm), especially when compared to touch typing on a physical keyboard, which ranges between 40 to 60 wpm [1, 4].

A 'switch' is an alternate method of selection that is used to produce input for an AAC system or technologies, such as a tablet. There are many types of 'switches', the most common type being one that is pressed by a finger, hand, head or other body part. However, for those with limited controlled movement, other alternatives must be considered. Often this is eye gaze, but eye gaze with dwell can be slow and frustrating. Thus, the use of an EMG switch for selection, which can detect bioelectrical activity with such little movement, can be invaluable. This study uses EMG signals made by the user as a switch. The switch can be attached in various ways, commonly via a band as seen below in *Figure 2*.

Figure 2: Render of a model using an eye gaze system with an EMG switch worn as a band on the wrist

The purpose of this prospective, crossover trial is to compare the speed, accuracy, and other qualitative factors of communication, using eye gaze with dwell and eye gaze with EMG switching for selection methods. This comparison will provide insight into how to provide fast and accurate communication to AAC users.

2. Method

Participants

The 30 participants chosen were all able-bodied individuals with a good understanding of the English language, without glasses, and novice users of eye gaze dwell and EMG switching. Despite the technology being suited for the disabled population, able-bodied individuals were used for the study to avoid introducing too many variables, skewing the data, and making comparison on the technical aspects of the technology inconclusive. The participants age ranged from 20 to 26 (with a mean of 21.33 and a standard deviation of 1.21.)

Equipment

Eye Gaze DUO® (Eye Gaze with Dwell) *from Control Bionics*

A TM5 Mini Camera® from EyeTechDs, integrated into a 12-inch Windows tablet running Windows 10 Home® (with 2736 x 1824 resolution) was used for the tracking of the participants' gaze. Eye gaze was used to access a QWERTY keyboard on the latest technology of Grid 3® software by Smartbox® Assistive Technology. The letter keys were of 2.2 x 2.5 cm and the space bar was 2.5 x 17.5 cm.

NeuroNode Trilogy® (Eye Gaze with EMG switching) *from Control Bionics*

The NeuroNode Trilogy device contains the same hardware and software as mentioned above in the Eye Gaze DUO; however, this system also has an EMG switch, named the "NeuroNode," as seen below in *Figure 3.*

Figure 3: The NeuroNode, the EMG switch used for this research

Software set up

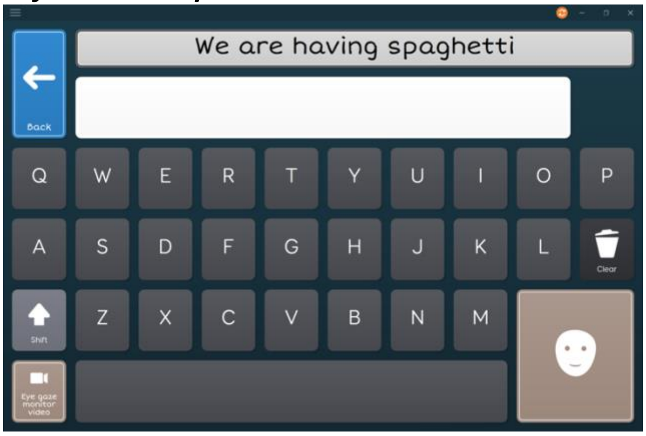

Figure 4: Screenshot of the keyboard layout for the test, with the example phrase of "We are having spaghetti"

The stimulus phrase was located at the top of the screen, just above the input section, for easy referencing as illustrated in *Figure 4*. There were eye gaze monitor video buttons which allowed the user to see whether they were in the appropriate position to be recognised by the camera, thus ensuring optimal eye movement detection. The stimulus phrase, the buttons 'Back', 'Clear', and the eye gaze video keys were restricted to 'touch' only, meaning the participant could not activate them through their access method.

When using the eye gaze system with dwell, there was a dwelling circle animation around the letter being chosen to indicate the progression of the dwell, with a full circle representing activation alongside a clicking sound to assist in providing audio feedback. The clicking sound was present in both systems, however, there was no circle animation in the eye gaze with EMG switching system since there was no dwell time involved.

Figure 5: NeuroNode set up on forearm to measure index finger flexion

For the purpose of the trial, the NeuroNode was placed on the participant's wrist as seen in *Figure 5* to measure and use the muscle activity in the *flexor digitorum superficialis*, which is the muscle in the forearm that is activated to flex the index finger. This includes the NeuroNode on the NeuroBand attachment, which is held together by an elastic strap, so that the NeuroNode sits securely in the intended position, with the centre of the NeuroNode on top of the muscle. This allows the NeuroNode to be used as the EMG switch for this study.

Procedure

The sessions began with a quick briefing on the task instructions and a breakdown of how the session would proceed, including a full explanation of how the eye gaze and NeuroNode components operate. The task instructions to the participants were to type the phrases as quickly as possible without making any errors.

All participants used the same 2 sets of 6 phrases. Half of the participants started using the eye gaze with dwell first and then used eye gaze with EMG switching. The other half did the reverse, starting on eye gaze with EMG switching, followed by eye gaze with dwell to ensure the results were unbiased and fair. There was a 5-minute break in-between each set for the participant, to reduce the risk of fatigue. The camera was then calibrated before each set, using 16 targets on normal speed, with the participant sitting 50cm away from the screen. The calibration scores were recorded, and all targets had to be green (score of 75/100 or higher) to ensure the system was calibrated to an adequate level. If the participant could not record all green scores after 5 attempts, they were not able to participate in the study and the session was concluded. If the camera needed to be recalibrated, it would be ideally done in-between phrases. If it had to be redone during a phrase, the phrase was then repeated after the camera was recalibrated. The NeuroNode was calibrated in the official way recommended by Control Bionics.

The 12 phrases were taken from the *MacKenzie and Soukoreff* [5] phrase set, all performed in the same order under each condition and can be found under *Appendix A*. The phrases did not contain any punctuation or capital letters apart from the first letter which was auto capitalised. Additionally, predictive text was not used to avoid skewing of the data. The participants were given two non-recorded warm up phrases to familiarise themselves with each access method.

The participants were instructed to memorize the phrase before attempting to write it as quickly and accurately as possible. When an error was made, they were instructed to move on without correcting it. However, if the user skipped or wrote a completely different word, the phrase was repeated as this would skew the data significantly in a manner which was not directly related to the access method itself. The phrases were shown on the home screen as seen in *Figure 6* and read out by the study supervisor.

Figure 6: Phrase menu for set 1

The phrase was then selected, and the participant would begin. Once the participant began to dwell (for eye gaze with dwell) or make the movement to activate the EMG switch (for eye gaze with EMG switching) on the first letter, the supervisor started the stopwatch, ending it once the last letter in the sentence had been activated. The supervisor recorded the time of the phrase as well as any errors that were made. For the purposes of this study, the dwell time was set to 1000ms. This time was chosen because the majority of current Control Bionics users dwell time is between 800-1000ms, with some users as high as 1200ms, thus 1000ms was chosen as it represents a majority of actual users.

The session ended with a Participant Survey to collect qualitative data on their experiences using each system in areas such as enjoyment and fatigue, and any preferences or additional notes they had. Each session went for approximately an hour and 15 minutes.

3. Results

There was a pool of 30 participants that typed a total of 720 recorded phrases. The results tested for speed in wpm, accuracy derived from the 'minimum string distance' (MSD) error rate, and a survey with answers on a 1-7 Likert scale.

Speed

A word was considered as 5 characters including spaces, which is consistent with other published studies [3, 6]. The grand mean speed of the eye gaze with dwell and the eye gaze with EMG switching systems were 6.83 wpm and 10.03 wpm, respectively. This returns a **47% increase in speed** when using the eye gaze with EMG switching as opposed to the original eye gaze methods, as seen below in *Figure 7*.

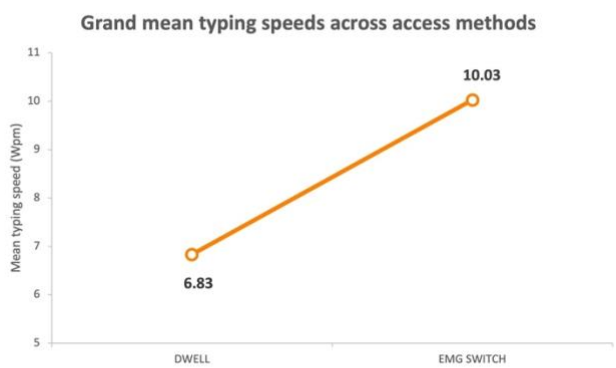

Figure 7: The grand mean of the typing speed for both systems

The order of the access methods which the participant performs has an impact on the typing speed as seen in *Table 1* below, especially for the eye gaze with EMG switching method.

Table 1: *Grand mean speeds in words per minute for different conditions*

Selection Mode	Eye Gaze with Dwell WPM	Eye Gaze with EMG Switch WPM
Dwell first	6.75	10.60
EMG first	6.92	9.46
Overall Average	6.83	10.03

Accuracy

The 'minimum string distance' (MSD) error measures the minimum number of edits (insertions, deletions, or substitutions) required to transform one string into another, which in this case is the transcribed text into stimulus text. For example, the presented text of "Three two one zero blast off" transcribed as "Therwe two one zero blast off" would have a MSD of 2. This is from the deletion of the first 'e' and substitution of the 'w' to 'e' in 'Therwe' to make 'Three'. The grand mean of the MSD error instances for eye gaze with dwell and eye gaze with EMG switching systems were 3.27 and 4.03 over the 344 characters. The percentage of successful character presses returned accuracy rates of **99.05% and 98.83%** for eye gaze with dwell and eye gaze with EMG switching, respectively as shown in the graphical representation in *Figure 8.*

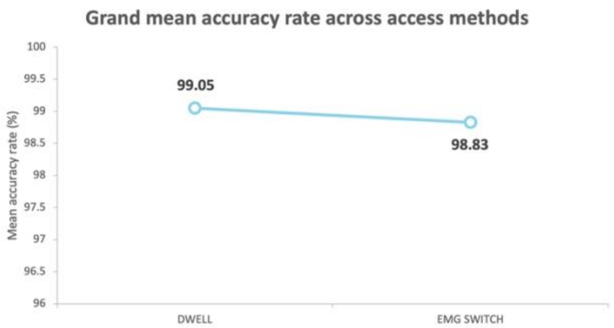

Figure 8: The grand mean of accuracy rates for both systems

Participant Survey

The results of the survey are below in *Table 2.*

Table 2: Survey results and insights

STATEMENT	AVERAGES		INSIGHT
	Dwell	EMG Switch	
I liked writing with this system	5.1	5.7	Participants liked writing with the eye gaze with EMG switch system **8.57%** more.
I found this system stressful to use	3.13	2.67	Participants found eye gaze with EMG switch system **6.57%** less stressful to use.
I found that this system required lots of concentration to use	5.73	4.23	Participants found that eye gaze with EMG switch system required **21.43%** less concentration to use.
I feel like my use of this system improved with practice	6.27	6.23	Participants found that proficiency in both systems improved with practice.
I would find this system tiring to use over an extended period of time	6.03	4.13	Participants believed the eye gaze with EMG switch system would be **27.14%** less tiring to use over extended periods of time.

I could operate this system with ease	5.13	5.7	Participants found eye gaze with EMG switch system **8.14%** easier to use.
I found this system natural to use	4.73	5.4	Participants found eye gaze with EMG switch system **9.57%** more natural to use.
I experienced mental overload when using this system	3.67	2.8	Participants experienced **12.43%** less mental overload when using eye gaze with EMG switch system.
I feel like I was typing quickly with this system	4.2	5.73	Participants felt like they were typing **21.86%** quicker when using eye gaze with EMG switch system.
I prefer writing with '_______' system	7	23	**76.67%** of participants preferred writing with eye gaze with EMG switch system.

4. Discussion

Speed

The results presented the eye gaze with EMG switch being a quicker typing system than traditional eye gaze with dwell, evident in the 47% increase in typing speed. While it is likely that eye gaze with dwell typing speed would increase over time [3, 7], it is likely that eye gaze with EMG switching speed would also increase with further practice and familiarity with the system. The faster eye gaze with EMG switching typing can be attributed to the absence of dwell time required to select a button, as once the letter has been cognitively processed and located, the participant can immediately activate it with the EMG switch, then move onto the next letter in the stimulus phrase.

Table 1 presents the various typing speeds depending on the order of the access methods; it shows that the first method was done more slowly than the overall average and the second method was done faster than the overall average. This suggests that the participants are likely to be improving with the eye gaze technology regardless of the dwell or EMG switch selection as the same eye gaze technique is being used to locate the letters. As such, this supports the idea that further practice and usage of the eye gaze will correlate to faster times. Interestingly, the larger variation in speeds for the eye gaze with EMG switching suggests that there is steeper initial learning curve when learning to use this system.

Further, the rising accuracy and effectiveness of predictive text may allow for phrases to be transcribed faster than the 6.83 and 10.03 wpm speeds produced in this research. The typing speed metric of words per minute defined a word as five key presses. Thus, if a predictive text system was enabled, the user could theoretically type faster than the previously mentioned speeds. For example, if the user was attempting to type a 9-key word like 'alligator' and after typing 'alli-', the full word is suggested and pressed as a predicted text, it would be a total of five clicks not nine. There is insufficient research on eye gaze predictive text typing to reference when considering the impact it would have for the user.

Accuracy

The accuracy values show eye gaze with dwell being 0.22% more accurate than eye gaze with EMG switching. This difference is from 99.05% on eye gaze with dwell, to 98.83% on the eye gaze with EMG switching, thus there is an insignificant difference in accuracy between the two methods. Due to the same eye camera being used on both systems, differences in accuracy would be related to the ability of the user to concentrate and execute their phrases.

Participant Survey

The survey results generally reflected that the participants found the eye gaze with EMG switch method to have better qualities than its counterpart. A difference of importance came in the mental overload question, where the participants found they experienced more mental overload when using the eye gaze dwell method as opposed to the eye gaze with EMG switch. The definition of mental overload was explained as '*an experience which is so cognitively and perceptually stimulating that it taxes or exceeds the individual's capacity to process existing or incoming information*' [8]. The survey results reflect a reduced cognitive strain required by the user to operate the eye gaze with EMG switching. Another insight of interest was regarding the eye gaze with EMG switch being less fatiguing over extended use. This question had to be asked theoretically as typing 4 sets of 6 phrases with five-minute breaks in-between each set would likely would not be enough time to experience expected levels of fatigue that would occur with regular use of up to 10 hours per day for some individuals.

5. Conclusion

By removing the dwell time and replacing it with the EMG switch, there is no longer a need to dwell to activate a key, eliminating a large portion of the typing time. This produces a more natural experience, similar to that of a computer mouse, allowing one action to choose the key and another action to instantly select the key. This advantage is reflected in the 47% increase in typing speed when using the eye gaze with EMG switch.

This ease to select items on the eye gaze with EMG switching shows no significant impact on the accuracy when compared with traditional eye gaze with dwell. Thus, there is no consequence for the increased speed that can be achieved when using eye gaze with EMG switching.

The pressure to avoid unintentional selection whilst looking at keys appears to be a source of the mental overload when using eye gaze with dwell. The implementation of an EMG switch eliminates this problem.

The limitations of this study are the lack of exposure towards the learning and improvement as a user becomes more experienced and familiar with the system. Additionally, the use of predictive text was not explored which may have produced faster typing speeds and is commonly used by AAC users. Lastly, only using able-bodied participants for access methods used by people with mental or physical impairments, or both. These elements could be considered for future investigations if further data was required on the true maximum speed and potential of the systems. Overall, the results do show that the use of the eye gaze with EMG switching produces faster and less fatiguing typing than traditional eye gaze dwell methods. The conclusions from this research provide an insight on the optimal access method for those with motor diseases and communication difficulties.

Acknowledgments

I would like to thank the individuals who participated in the research for giving their time to do so. Additionally, I would like Control Bionics for providing the resources required to execute the research.

References

[1] M. E. Mott, S. Williams, J. O. Wobbrock and M. R. Morris, “Improving Dwell-Based Gaze Typing with Dynamic, Cascading Dwell Times,” *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems,* pp. 2558-2570, 2017.

[2] K. S. Karn, “Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises,” *Mind,* vol. 2, no. 3, pp. 573-605, 2003.

[3] P. Majaranta, U.-K. Ahola and O. Špakov, “Fast Gaze Typing with an Adjustable Dwell Time,” *Proceedings of the 27th international Conference on Human Factors in Computing Systems (CHI '09),* pp. 357-360, 2009.

[4] P. Majaranta and K.-J. Räihä, “Text entry by gaze: Utilizing eye-tracking.,” in *Text Entry Systems: Mobility, Accessibility, Universality*, San Francisco, Morgan Kaufmann, 2007, pp. 175-187.

[5] I. S. MacKenzie and R. W. Soukoreff, “Phrase Sets for Evaluating Text Entry Techniques,” *In CHI '03 Extended Abstracts on Human Factors in Computing Systems (CHI EA '03),* pp. 754-755, 2003.

[6] P. O. Kristensson and K. Vertanen, “The Potential of Dwell-Free Eye-Typing for Fast Assistive Gaze Communication,” *ETRA '12: Proceedings of the Symposium on Eye Tracking Research and Applications,* vol. 2012, pp. 241-244, 2012.

[7] O. Tuisku, P. Majaranta, P. Isokoski and K.-J. Räihä, “Now Dasher! Dash Away! Longitudinal Study of Fast Text Entry by Eye Gaze,” *ETRA '08: Proceedings of the 2008 symposium on Eye tracking research & applications,* pp. 19-26, 2008.

[8] “overload,” American Psychology Association, 2020. [Online]. Available: https://dictionary.apa.org/overload. [Accessed 15 October 2021].

Appendix A: Phrases List

Warm up phrases

- We are having spaghetti
- Time to go shopping

Set 1

- My watch fell in the water
- Prevailing wind from the east
- Never too rich and never too thin
- Breathing is difficult
- Physics and chemistry are hard
- Elections bring out the best

Set 2

- A problem with the engine
- Elephants are afraid of mice
- My favourite place to visit
- My favourite subject is psychology
- Circumstances are unacceptable
- Watch out for low flying objects