Quantitative and Comparative Assessment of Learning in a Tongue-Operated Computer Input Device

Behnaz Yousefi[Student Member, IEEE],
GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA

Xueliang Huo[Student Member, IEEE],
GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA

Emir Veledar[Senior Member, IEEE], and
Cardiology Division, Emory University School of Medicine, Atlanta, GA 30332 USA

Maysam Ghovanloo[Senior Member, IEEE]
GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA

Behnaz Yousefi: byousefi@gatech.edu; Xueliang Huo: xhuo@gatech.edu; Emir Veledar: eveleda@emory.edu; Maysam Ghovanloo: mgh@gatech.edu

Abstract

Tongue drive system (TDS) is a wireless, wearable assistive technology that enables individuals with severe motor impairments to access computers, drive wheelchairs, and control their environments using tongue motion. In this paper, we have evaluated the TDS performance as a computer input device in four tasks, commonly known as horizontal, vertical, center-out, and multidirectional rapid tapping, based on Fitts' law and ISO9241-9 Standard. Nine able-bodied subjects, who already had tongue piercing, participated in this trial over five sessions during 5 weeks, allowing us to study the TDS learning process and its current limiting factors. Subjects wore tongue rings made of titanium in the form of a barbell with a small rare-earth magnetic tracer hermetically sealed inside the upper ball. Participants performed the same tasks with a mouse (only in the first session) as a reference as well as a standard keypad for benchmarking. Six performance measures were considered, including throughput, error rate, and reaction time, all of these improved significantly from the first to the last session, and some of these plateaued over the course of the experiment. The comparison between tongue–TDS versus index-finger–keypad provides valuable insights into tongue human factors, which can lead the way in improving the usability of the TDS and similar tongue-operated assistive technologies.

Index Terms

Assistive technologies (ATs); environmental access; Fitts' law; human–computer interaction (HCI); severe disabilities; Tongue Drive System (TDS)

I. Introduction

Complete recovery from neurological injuries and diseases that lead to severe movement disabilities is expected to arise from research in regenerative and repair medicine at the molecular and cellular levels [1]. However, safe and widely accepted methods might be decades away and in the meantime there is a need for more immediate solutions [2]. Among the most important issues that degrade the quality of life of severely disabled individuals is
the loss of independence [3]. Substituting a portion of their lost abilities or augmenting the remaining ones effectively with advanced technologies is the ultimate goal of many biomedical engineers. They endeavor to either tap into the human nervous system by developing advanced neuroprosthetic devices or utilize any remaining abilities of the paralyzed individuals with modern assistive technologies (ATs) [4], [5]. Even though the current solutions fall short of natural human abilities, regaining even a partial independence in tasks such as computer access, wheeled mobility, and environmental control can tremendously help individuals with severe disabilities to catch up with today’s fast paced lifestyle, which pivots on information technology and communications.

According to [6], ATs for the mobility impaired can be categorized as: 1) ATs that are based on physiological signals, such as electroencephalogram (EEG), electromyogram (EMG), and electrooculogram (EOG) [4], [7]–[12]; 2) ATs that track movements of body parts by various sensors, such as head and eye trackers [13]–[20]; 3) ATs based on voice commands, such as speech recognition and nonverbal vocalization software [21], [22]; and 4) the mechanically actuated ATs, such as sip-and-puff or mouth stick [23], [24]. There are also hybrid ATs that are the combinations of more than one category [9], [14].

Tongue drive system (TDS), which can be placed in the second group, consists of a small permanent magnetic tracer fixed on the tongue with adhesives or piercing, a headset with an array of 3-axial magnetic sensors to detect the changes in the magnetic field generated by the tracer, a wireless link between the headset and a transceiver on a computer, or a smartphone to transfer the magnetic sensor data, and a sensor signal processing (SSP) algorithm, which recognizes the position of the magnetic tracer and, hence, the position of the tongue (see Fig. 1) [25]–[28].

The TDS allows its users to associate a specific set of tongue gestures, which they can remember and repeat consistently, with a set of commands that can be defined for the equipment of interest, such as computer, wheelchair, phone, TV, etc. The current TDS prototype can offer six simultaneously available commands, including four directional commands (LEFT, RIGHT, UP, and DOWN) and two selection commands (LEFT-SELECT and RIGHT-SELECT). When using the TDS for computer access, for instance, the directional commands move the cursor on the screen in four cardinal directions and the selection commands emulate the mouse single/double-clicking functions. The tongue resting position is detected as the NEUTRAL command.

The TDS is designed to serve individuals with disabilities in their upper limbs due to neurological injuries or diseases, who have maintained their voluntary tongue motions. Individuals with high-level spinal cord injury (SCI) at C5 and above, amyotrophic lateral sclerosis (ALS), mild cerebral palsy, and two-arm amputees are among the TDS target population. A brief description of the potential benefits and limitations of using tongue as a manipulative appendage along with a list of TDS characteristic features in comparison with other ATs have been previously covered in [25], [26], and [28].

II. Background and Scope

Research shows that maximum comfort and satisfaction of the AT users are highly correlated with the AT performance, provided that it matches the person in key areas, such as functional needs, lifestyle, environment, economic status, support network, etc. [5], [29]. Thus, every new AT needs to be quantitatively and comparatively assessed according to widely accepted performance measures to inform the stakeholders, such as clinicians, rehabilitation professionals, caregivers, payers, and potential end users on one hand, and to guide future improvements of that device and other similar technologies on the other hand.
Considering the wide range of disabilities and the large variety of ATs that are currently available, a direct comparison between a new AT and those in the market requires recruiting a large number of individuals with disabilities. An extensive set of trials should be designed in which every subject is allowed to learn how to use the new AT, as well as its alternatives, and, then, perform a variety of tasks that resemble everyday-life activities with every AT for a certain period of time. Even though conducting such an experiment is not impossible, in terms of cost and logistics, it is beyond the reach for many academic groups. What makes such a direct comparison even more complicated is the fact that individuals with severe disabilities often use one or more ATs on a regular basis, and as a result, there will always be a bias toward the AT(s) that they are most familiar with.

An alternative approach, at least for the ATs that are meant to be used as computer input devices, is to benchmark them against the most popular input devices (keyboard/mouse) that are well characterized by the human–computer interaction (HCI) community as the gold standards, while adopting widely accepted measures and performance metrics. Now, if the other ATs are also evaluated in the same fashion, it would be possible to indirectly compare the new AT with others. One limitation of this approach is that the test subjects should be healthy individuals, who are not the intended end users of the new AT. A second limitation would be the subjects' familiarity with the standard devices, particularly, if they are recruited from the college student population. On the other hand, unlike the direct approach, there will be a certain level of homogeneity among subjects in terms of their ability in using standard devices and their naivety toward the new AT.

In our earlier studies, we measured the TDS information transfer rate, a widely accepted performance measure in brain–computer interfacing, as well as the ability to move the cursor through an on-screen maze. Both able-bodied and SCI subjects used the TDS only in one session by attaching the TDS magnetic tracer on their tongues using dental adhesives [25], [28]. Here is what we are presenting for the first time in this study.

1. Evaluating the TDS performance for cursor control using ISO9241-9 Standard tasks for pointing and selecting along with a comprehensive set of measures [30]. A key parameter in this standard is throughput, measured in bits/s as an indicator of the amount of information that users can deliver to a computer through the device under test. ISO9241-9, which is based on the well-known Fitts' law, has been widely adopted by the HCI community for evaluating conventional nonkeyboard input devices, such as mouse, trackball, and touchpad [31], as well as ATs, such as eye trackers [19], head trackers [6], [10], and voice-activated software [22]. Fitts' law, proposed in 1954, models human motor behavior by stating that any human motor action in performing a certain task conveys a finite amount of information that is characterized by a tradeoff between speed and accuracy [32]. Fitts' law is linked to Shannon's theorem in information theory, which in turn describes the information capacity limits of a communication channel [33]. ISO9241-9 Standard shows how to calculate the throughput in certain tasks of rapid movements over on-screen targets of different sizes and distances with the purpose of emulating human interactions with GUIs.

2. Evaluating the TDS performance over five sessions during 5 weeks, as opposed to a single session, in order to observe the learning process, which is an important factor in the acceptability and adoption of a new AT.

3. Evaluating the TDS performance by nine able-bodied subjects who had already received tongue piercing. The magnetic tracer was embedded inside the upper ball of barbell-shaped titanium tongue studs and worn by the subjects throughout the study, as opposed to being temporarily glued on their tongues.
To compare the tongue–TDS–computer performance with that of the index-finger–keypad–
computer, similar cursor-control tasks were performed with both devices. Moreover, to
validate our experimental methods and data analyses, the study included performing all tasks
with a standard optical mouse, for which the range of performance measures has been well
established in the literature [31].

III. Methodology

Subjects performed seven tasks (with the number of employed TDS commands in
parentheses): horizontal (2) and vertical (2) unidirectional tapping, on-screen maze
navigation (4), center-out tapping (4), issuing timed random commands (6), multidirectional
tapping (6) and driving a wheelchair in an obstacle course (4). Here, we report on the
tapping tasks, i.e., uni- and multidirectional tapping and the center-out tapping, which are
among the most commonly used tasks on the basis of Fitts’ law [4], [9], [10], [14], [34]–
[38]. The rest of this trial and corresponding results have been reported elsewhere [39].

To facilitate learning, particularly, in the first session, these tasks were arranged from easy
to difficult in terms of the required number of TDS commands. The GUI was developed in
the LabVIEW environment and presented to the subjects sitting 1 m away from a 22” LCD
monitor with 1280 × 800 pixel resolution. The task window was a white 610 × 610 pixel
square in the center of the monitor, shown in Fig. 1(b). Within each task, the device order
(TDS, keypad, mouse) was randomized and its icon appeared on the right-hand side of the
screen. Each task with each device was performed in four rounds, the first of which was for
practice. Subjects were allowed to rest up to 1 min between consecutive rounds.

In tapping tasks, targets represent simplified versions of the icons, links, and buttons in
everyday usage of the GUIs, whose dimensions, locations, and orientations on the screen
often guide the design of such experimental setups. In our case, the choice of the task-screen
size, number of targets, their widths and distances, and the number of condition repetitions
were primarily selected based on similar instances in the literature [6], [10], [19], [22], [35]–
[38]. The total number of trials in each task was defined by limiting the duration of each
session to 3 h while maintaining enough statistical power to analyze each condition.

A. Tasks

1) Unidirectional Tapping—A pair of rectangular target bars with three different widths
\((W = 30, 61, \text{and } 122 \text{ pixels})\) and center-to-center distances \((D = 122, 244, \text{and } 488 \text{ pixels})\)
created a total of 9 \(D–W\) pair conditions (see Table I), which appeared one at a time on the
screen as shown in Fig. 2(a). Subjects were instructed to move the cursor as fast and as close
to the center of the bars as possible in back-and-forth movements. Nine conditions appeared
randomly, 3 per round, and each condition repeated 18 times for a total of 54 \((3 × 18)\) trials
per round. At the beginning of each round, the cursor was located at the center of one of the
target bars and subjects began that round by heading toward the other bar. While using the
TDS and keypad, subjects were required to reach the target bars and change the direction of
movement using \(\text{LEFT/RIGHT}\) and \(\text{UP/DOWN}\) commands in the horizontal and vertical
unidirectional tappings, respectively. When using mouse, they were also required to click on
the target bar.

2) Center-Out Tapping—Circular targets with three different diameters \((W = 30, 61, \text{and }
122 \text{ pixels})\) and three different distances from the center of the screen \((D = 61, 122, \text{and } 244
\text{ pixels})\) created a total of 6 \(D–W\) pair conditions (see Table I). They appeared randomly on
the screen one at a time along cardinal and ordinal directions (every \(45°\)), making 48 \((6 × 8)\)
trials per round, as shown in Fig. 2(b). Subjects were instructed to move the cursor as fast
and as close to the center of the targets as possible. Each trial began with the cursor at the
center and a new random target on the screen. Using the TDS and keypad, subjects were required to dwell for 560 ms, as suggested in [19], to select a target. With mouse they were asked to click on the target.

3) Multidirectional Tapping—Fifteen equally spaced circular targets with two different widths ($W = 57$ and 76 pixels) appeared around the circumference of a circle with two different diameters ($D = 305$ and 534 pixels), creating a total of 3 $D–W$ pair conditions (see Table II), as shown in Fig. 2(c). Targets turned blue one at a time in a sequence that rotated clockwise across the diameter of the larger circle [see the numbers in Fig. 2(c)], while subjects were instructed to move the cursor as fast and as close as possible to the center of the targets. Each round included three sequences, corresponding to one of three randomly selected $D–W$ conditions, while each sequence was one complete rotation of the circular targets (1–15), resulting in a total of 45 trials per round. While using the TDS and keypad, subjects were required to issue a SELECT command for 460 ms to select a target. The selection period was intended to reduce the GUI sensitivity to unintentional rapid selections. With mouse they had to click on the target.

B. Input Devices

1) Tongue Drive System—Fig. 1(a) shows the TDS prototype used in this study [28]. Subjects wore barbell-shaped titanium tongue studs with 12 or 14 gauge posts, which were 12 or 15 mm in length (Anatometal, Santa Cruz, CA). The upper M&M-shaped ball ($ø8$ mm x 3.5 mm), which was laser welded to the post, had a disk-shaped ($ø4.8$ mm x 1.5 mm) rare-earth permanent magnet (K&J Magnetics, Jamison, PA) embedded in it. The magnetic flux density on the upper surface of the M&M ball was $1194 \pm 83$ G on average. The lower ball tightly screwed on to the post with a large number of threads. A pair of 3-axial magnetoresistive sensors was held bilaterally symmetrical to the sagittal plane, near the subjects’ cheeks with a pair of goosenecks to allow for adjustments. A small control unit on top of the headset housed a pair of AAA batteries and a low-power microcontroller with build-in 2.4 GHz RF transceiver (CC2510, TI, Dallas, TX) sampling each sensor output at 50 Hz. All samples were packed in one data frame, and transmitted wirelessly to a laptop, equipped with a USB receiver dongle [28].

The SSP algorithm running on the PC used a $K$-nearest-neighbor (KNN) classifier to identify the incoming samples based on their features, which were extracted via principal components analysis from the data that was collected during a training stage [25]. The tongue positions for the six TDS commands plus the tongue resting position, which was considered the NEUTRAL command, are shown in Fig. 3(a). It should be noted that the cursor movement in each direction was unlatched, i.e., the subjects had to continue issuing a certain directional command to maintain the cursor motion in that direction. The speed of cursor movement was increased linearly from zero at a rate of 500 pixels/s$^2$ until it reached saturation level at 350 pixels/s. The cursor speed profile (acceleration rate and saturation level) was determined based on the comfort level of an experienced TDS user. Although in practice such parameters can be customized based on the users’ skill levels and preferences, we set these values to a level that challenged our naive subjects.

Reaching targets in oblique directions with four directional commands requires moving in horizontal/vertical segments. Since the cursor speed ramped up every time a new command was issued, the most efficient way of reaching such targets was to minimize the number of segments. Even then reaching oblique targets always took longer than reaching horizontal or vertical targets at a similar distance. To prevent subjects from adjusting the cursor position with a series of small moves and to prevent issuing wrong commands repeatedly, a time
limit was set in reaching a target, beyond which the trial was terminated and a new target appeared.

2) Keypad—Fig. 3(b) shows the keys on a standard keypad designated to the same commands used in the TDS in a way that they resembled their positions in the mouth. The keys were also chosen such that subjects have least difficulty locating them while looking at the GUI screen. Subjects were required to issue commands by pressing the keys using their right index finger, while comfortably laying their hand on the table. The keypad output was sampled at 50 Hz to be consistent with the TDS and the same cursor velocity profile was applied.

C. Performance Measures

The following performance measures were considered for the quantitative assessment of the TDS performance and its comparison with mouse and keypad [10], [31], [35].

1) Throughput—Throughput (TP) shows the amount of information that subjects delivered to the computer through each of the input devices. TP is defined as the ratio between the effective index of difficulty, $\text{ID}_e$, of targets with the same condition (same $D$–$W$ pair) to the average time it takes to reach them, $\text{MT}$:

$$\text{TP} = \frac{\text{ID}_e}{\text{MT}}$$  \hspace{1cm} (1)

where $\text{ID}_e$ of the target, measured in bits, is defined by Shannon’s formula as

$$\text{ID}_e = \log_2 \left( \frac{D_e}{W_e} + 1 \right).$$  \hspace{1cm} (2)

Here, $W_e = 4.133 \times \text{SD}_x$ is the effective width for each condition, and $\text{SD}_x$ is the standard deviation of $x$, which is the distance between where the subject points while reaching a target to the center of that target, when projected onto a straight line from the origin of the movement to the center of the target, known as the task axis [36], [37]. ‘$x$’ can be positive or negative when the subject overshoots or undershoots during a pointing task, respectively. The effective distance $D_e$ is defined as the mean of the distances of the pointed spots projected along the task axis over all the targets with the same condition. “MT” in (1) only includes the time when the cursor is moving, i.e., it neither includes the initiation delay time before the subject moves the cursor nor the dwelling or selection time for target selection. With the aforementioned definition, TP bears both the speed and accuracy of the subjects’ pointing performance [31], [35].

Tables I and II show the IDs for different $D$–$W$ pair conditions in our tapping tasks. ID values are derived from (2), where actual $D$'s and $W$'s are inserted rather than $D_e$'s and $W_e$'s. Although according to [30] and [31] subjects should be presented to a wide range of IDs, we limited this range to 1–4.11 bits, similar to [37], to match the expected usage of the TDS for selecting icons and buttons. The overall TP for each device, each session, and each task was computed as the grand mean of TPs of all subjects [31].

2) Error Rate—Error rate (ER) is the percentage of the taps outside the targets to the total number of taps for each task. While TP does not reflect whether the targets were eventually selected or not, the ER reveals the subjects’ accuracy in using the computer input device for
pointing and selecting the targets [31], [36]. We asked the subjects to focus on accuracy as opposed to speed if their ERs tend to be growing.

3) Task Completion Time—Task completion time (TCT) is the total time it took to complete each task, excluding the practice round and pauses in-between rounds. TCT includes the initial time before cursor movement, cursor movement time, and the times when cursor was stopped because no command had been issued or dwelling/selection commands were in progress.

4) Path Efficiency—Path efficiency (PE) is the ratio of the length of a direct path between the center and selected points in the center-out or between two consecutive selections in the multidirectional tasks to the actual path that the cursor has traversed, averaged over all three rounds of that task. PE is meant to indicate the overall subjects' control over the cursor movements.

5) Deviation From Optimum Number of Movements—To achieve lower TCT, subjects were instructed to minimize the number of movement segments (see Section III-B). From the center of the task window, half of the 48 targets in the center-out task were along the horizontal or vertical axes and the other half were along the diagonal axes, requiring at least 24 and 48 cursor movements to reach them, respectively [see Fig. 2(b)]. Hence, completing all three rounds required at least 216 segments in an optimal performance. Similarly, the minimum number of segments required in the multidirectional tapping task was 261 [see Fig. 2(c)]. Deviation from optimum number of movements (DONM) was defined as

\[
\text{DONM} = \left| \frac{\text{Total Number of Movement Segments}}{\text{Optimal Number of Movement Segments}} - 1 \right|.
\]

Although lower DONM does not necessarily mean a good performance, along with other performance measures it can further clarify the subjects' performance.

6) Reaction Time—Reaction time (RT) is the time interval between when a new target is shown on the screen and the initiation of the cursor movement in the center-out tapping task. It includes subjects' decision time to which direction to move, which tongue position or key to select, physical movement of the tongue or finger, and the system latency (including the GUI) in detecting and applying subjects' commands to cursor movements. The TDS and keypad latencies were measured 160 and 80 ms, respectively. RT for each device, session, and subject is the average of all reaction times over three rounds.

D. Protocol

The necessary approval was obtained from the Georgia Tech's institutional review board (IRB). Nine right-handed able-bodied subjects, four males and five females, aged 19–28 years, who had tongue piercing in the midline of their tongues, between the tip and frenulum, completed the trial out of 14 who were initially recruited. The subjects were all naive and did not have any prior experience with the TDS. Each subject was scheduled on a certain day of the week and allowed to have up to ±2 days variation. Two consecutive cancellations terminated the subjects' participation in the trial. On average, the first session, in which subjects became familiar with the TDS, its operation, and the tasks, took ∼5 h while the following sessions took ∼2.5 h.
In the first session, subjects were asked to swap their studs with the cold-sterilized TDS magnetic tongue studs and wear them throughout the 5-week trial. At the beginning of each session, subjects wore the TDS headset (see Fig. 1) and moved through calibration and pretraining steps as described in [25]. Following training, tasks were arranged from easy to difficult by increasing the number of TDS commands in three stages (2, 4, and 6 commands, plus NEUTRAL).

E. Data Analysis

Considering that the main input devices were TDS and keypad, our experiment was a 5 × 2 within-subject factorial design with two factors of session (5 levels) and device (2 levels). We focused on two key aspects of the current TDS: 1) learning curve and 2) performance limiting factors. The TDS learning curve was characterized with the initial performance, overall performance improvement, and whether performance reached a plateau during five sessions. TDS limiting factors were linked to the difference between the subjects' performance with the keypad and TDS, particularly in the fifth session. Each task in each session started with a practice round followed by three main rounds and the order of the TDS or keypad was randomized. Since tongue performance is independent of the index finger, to focus on the TDS learning curve, we used one-way repeated measures ANOVA, taking only the TDS performance into account. We also employed Helmert contrast to find non-significance between the performance of each session and the remaining sessions. We have defined reaching a plateau as a nonsignificant change in performance in at least the last two sessions with \( p > 0.05 \). To investigate the TDS limiting factors, we compared the TDS and keypad in the fifth session by applying paired \( t \)-tests.

IV. Experimental Results

Figs. 4–6 show the performance measures with their 95% confidence intervals of unidirectional, center-out, and multidirectional tapping tasks, respectively, for the three input devices, i.e., mouse (as a reference), keypad, and TDS, throughout five sessions. A summary of the results along with key statistical outcomes has been included in Table III. Third column shows the subjects' performance measures with mouse, which was only measured in the first session to generate a reference point. All mouse TPs are within the generally accepted range of 3.7–4.9 bits/s [31], which validates our methodology, GUI functionality, and data analysis.

Fourth and sixth columns in Table III show the TDS performance measures in the first and fifth sessions, respectively, which are used in paired \( t \)-tests with the resulted \( p \) values in seventh column indicating their significant difference (when \( p < 0.05 \)). These comparisons show that subjects' performances have been significantly improved in 14 out of 17 performance measures over five sessions. In the remaining three measures, the effect of session was to improve the average values (a reduction in these cases), and also reduce their variability, although their corresponding datasets are not significantly different. Similarly, sixth and eighth columns can be used to compare the TDS versus keypad performance measures in the fifth session. We have selected the fifth session for this purpose, because the subjects' acquired skills in using both devices were at their peaks within the span of this study. It should, however, be noted that while all subjects were naive with respect to the TDS, they had considerable prior experience with the keypad, and some used it on a daily basis. Therefore, as expected, the comparison showed that keypad performances were significantly better than the TDS in the fifth session in all but one performance measure. More interesting was the fact that the subjects' TDS versus keypad performance had converged over the course of five sessions, which is quite evident in the key measures of TP and ER in Figs. 4(a) and 6(a) and Figs. 4(b) and 6(b), respectively. Thus, it is not unreasonable to predict that if the number of sessions were higher, i.e., using the TDS on a
daily basis over an extended period of time, the TDS performance could become even closer to that of the keypad. This would not only depend on the subjects’ skills but also on the biomechanics and neuromuscular control of the index finger versus tongue motions, which are out of the scope of this paper [41].

To explore whether the subjects’ performances with the TDS had reached a plateau, we have indicated, in the fifth column of Table III, the level and the number of session (in parentheses) after which there were no statistical significance between the performance values \((p < 0.05)\). Despite this statistical result, all 14 plateaued TDS performance measures had improved with respect to their corresponding values in the plateau session. Most of those measures also showed reduction in the standard deviation, suggesting that subjects’ performances became more consistent over five sessions.

Fig. 7 shows sample cursor trajectory for one of the subjects in the center-out task with the TDS, keypad, and mouse in the first and fifth sessions. Improvement in the subject’s performance with the TDS in the form of more efficient traces and denser selection spots is evident both from Fig. 7(a) and (b) and performance values in the caption. It is also clear from Fig. 7(b) and (c) that the subject’s TDS performance became closer to the keypad in five sessions. Nonetheless, both the TDS and keypad, which are switch-based devices, therefore, confined to cursor motion in the cardinal directions with predefined speed profiles, are still inferior to mouse, which enjoys 360° directionality and proportional speed control [see Fig. 7(d)].

V. Discussion

Our main objective in this study was threefold: first, benchmarking the TDS performance by able-bodied subjects with respect to the gold standards, i.e., keypad and mouse, while adopting measures that are widely accepted by the HCI community; second, observing the learning process in using the TDS for computer access over five sessions, which included exploring the initial performance of able-bodied naive subjects, performance improvements, and the possibility of reaching a plateau; and third, revealing the limiting factors in the current TDS prototype by comparing it with another switch-based computer input device, with which the subjects were quite familiar, the keypad, operated by one of the most dexterous body parts, the index finger, for the exact same tasks, and number of discrete commands.

Fig. 8(a)–(c) shows the parameters affecting TP, which is the most comprehensive performance measure in this study, for the TDS and keypad in the fifth session of the center-out task. According to (1) and (2), these are the effective distance \(D_e\), effective width \(W_e\), and average movement time \(MT\) for each \(D–W\) pair. It is evident from Fig. 8(b) and (c) that the TDS MTs and \(W_e\)'s are higher than that of the keypad, resulting in lower TP. Higher MTs are expected considering higher DONM and lower PE in the TDS (see Fig. 5(d) and (e) and Table III), which indicates the subjects’ less control over the cursor movements using tongue–TDS versus index-finger–keypad. Although we could not find any direct comparison in the literature between the movement speed and accuracy of the tongue versus index finger, we can relate this outcome to the subjects' benefiting from visual and more pronounced tactile feedbacks when selecting and pressing a button down and releasing it versus touching a specific tooth with their tongues and returning to the resting position. Higher \(W_e\)'s in the TDS is the result of the selection spots being more scattered than the keypad, indicating less control. \(D_e\)'s in Fig. 8(a), however, were quite similar to the keypad and close to the actual \(D\) values (see Table I).
Using the cursor speed profile for both the TDS and keypad (see Section III-B), we calculated the minimum achievable MTs with a perfect performance for all $D-W$ pairs in the unidirectional tasks (i.e., center-to-center taps on targets with no stops in between) and overlapped them with the fifth session average MTs using thick solid lines in Fig. 8(d). It can be seen that for both the TDS and keypad, subjects have reached these minimums, which are dictated by the speed profile. Hence, the early plateaus in the unidirectional performance measures of TP or TCT (see Table III) are likely to be the result of cursor speed profile as opposed to either the subjects’ or the devices’ performance limits. Thus, to improve performance in switch-based controllers, it is important to optimize the cursor speed profile such that it does not become the bottleneck in achieving higher performance levels.

Considering that diagonal targets in the center-out task needed at least one stop to change the movement direction even with perfect performance, a $113–528$-ms delay, according to [4], was added to the minimum MTs to account for the subjects’ motor reaction time, plus a command latency of $120$ ms or $60$ ms for the TDS or keypad, respectively. The resulting optimal ranges have been depicted in Fig. 8(c) with hatched boxes over the measurement results. It can be seen that the higher the ID of each $D-W$ pair (in each column), the larger is the difference between optimal and measured MTs.

Fig. 8(c) and (d) shows that MTs in both the TDS and keypad have strong correlation with $D$ [$MT(W3, D3) > MT(W2, D2) > MT(W1, D1)$, where $D3 > D2 > D1$]. On the other hand, Fitts’ law predicts the MTs of $(W1, D1)$, $(W2, D2)$, and $(W3, D3)$ to be similar since they have the same IDs (see Table I) [30], [31]. The reason is using switch-based devices, for which the distance that the cursor traverses depends on the time duration of the command being issued. In other words, the further the target, the more time it takes for the cursor to reach it. Therefore, even with ideal performance, MT mainly depends on the $D$ of that target. On the other hand, in devices such as mouse that are proportional, MT depends on the speed and acceleration pattern of the hand movement, which, in turn, depends on both the target $W$ and $D$, and, consequently, on its ID [38], [40]. Although switch-based devices do not provide their users with the freedom of controlling the cursor speed, acceleration, and motion in any direction, the efficiency of cursor movements in terms of command timing, movement path, and number of movement segments are still under user control. This is why there are also strong correlations between MT and ID$_e$ for the TDS ($R^2 = 0.92$) and keypad ($R^2 = 0.98$), as shown in Fig. 9, for the fifth session of the center-out task.

In all tapping tasks, outside hits, which passed the targets (overshoot), were more than those, which did not reach them (undershoot). Also, in the center-out task, the number of outside hits for the ordinal targets was more than that on the cardinal targets. Because reaching diagonal targets needed at least two movements, which required two times the pointing precision. In all these cases, the number of outside hits with the TDS was more than that of the keypad, leading to a higher ER for the TDS.

The tongue anatomy and the biomechanics of its motion inside the oral cavity are important factors in the TDS users’ performance, which were not considered in this study [42]. Users’ performance also depends on the optimal position of the attachment of the magnetic tracer on the tongue. From a pure engineering perspective, the most agile position on the tongue that has the widest range of movements within the oral space is its tip, hence, the best position for the TDS magnet attachment. The further the magnetic tracer from the tongue tip toward posterior positions, the lower the range of motion, the lower the signal-to-noise ratio, and the worse the accuracy of the SSP algorithm in discriminating between adjacent TDS commands. However, there are also several other factors that identify the most appropriate position for the placement of the magnetic tracer, particularly, in the form of tongue...
piercing, such as the user comfort, oral safety, and risk of interference with speech and mastication [43].

In our previous studies, we were temporarily attaching the magnetic tracer \( \sim 1 \) cm away from the tongue tip using tissue adhesives [25], [26]. In this study, however, we had no control over the subjects’ tongue piercing position, which turned out to be \( 2.15 \pm 0.35 \) cm from the tip of their tongues [see Fig. 1(a)]. This obviously reduced the subjects’ agility and control in moving the magnetic tracer in their mouth. To this, we should add the small but important relative movements of the magnetic tongue stud with respect to the tongue, which can become more prominent if the length of the post is not properly selected with respect to the user’s tongue thickness. We hypothesize that the improvements in the subjects’ performance over time is not only because of learning how to control the computer cursor using the TDS, but also learning how to effectively manage the upper magnetic ball of the tongue stud in the mouth toward the TDS command positions [see Fig. 3(a)]. The information collected in this study, however, is not sufficient to discriminate between these two effects.

One limitation of this study, as mentioned in Section II, was recruiting healthy subjects, who are not the intended TDS end users, identified in Section I. Nonetheless, since the TDS solely relies on the users’ voluntary tongue motion and cognitive capacity, which are not compromised by SCI, upper-limb amputation, or many neurological insults, we expect similar results at least from this cohort of potential users.

VI. Conclusion

We have evaluated the TDS as a new tongue-operated computer input device in several standard tapping tasks via multiple widely accepted performance measures, such as throughput, over five sessions during 5 weeks, while nine able-bodied subjects, who already had tongue piercing, wore the TDS magnetic tongue studs over the course of this study. The experiment also included performing the same tasks with mouse as well as a standard keypad pressed by the right index finger for comparison with the TDS. Nearly all TDS performance measures improved significantly from the first to the last session and some reached plateau. Detailed comparison between subjects’ performance with the TDS and keypad provided valuable insights into the human factors of the tongue motion and a few limiting factors of the current TDS, which can lead the way in improving future revisions of the TDS as well as other similar devices that can also be used as ATs.

Acknowledgments

The authors would like to thank Dr. A. Laumann and her team at the Department of Dermatology, Northwestern University, Evanston, IL, for their constructive comments.

This work was supported in part by the National Institute of Biomedical Imaging and Bioengineering under Grant IRC1EB010915 and the National Science Foundation Award CBET-0828882 and Award IIS-0803184.

References


Biographies

Behnaz Yousefi (S’09) received the B.S. and M.S. degrees in electrical engineering from Khajeh Nasir University of Technology, Tehran, Iran, and Sharif University of Technology, Tehran, in 2004 and 2006, respectively. She is currently working toward the Ph.D. degree at the W.H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA.

She worked as an RF-Microwave Engineer for 3 years.
Xueliang Huo (S’07) was born in 1981. He received the B.S. and M.S. degrees in mechanical engineering (instrument science and technology) from Tsinghua University, Beijing, China, in 2002 and 2005, respectively. He is currently working toward the Ph.D. degree from the GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta.

His research interests include low power circuit and system design for biomedical applications, brain–computer interfacing, and assistive technologies.

Emir Veledar was born in 1954. He received the B.S. degree in economics from the University “Dzemal Bijedic,” Mostar, Bosnia, in 1976, the M.Sc. degree in statistics from the Institute of Economic Sciences, Belgrade, Serbia, in 1985, and the Ph.D. degree in statistics from the University “Dzemal Bijedic” in 1990. His Ph.D. research was focused on applying fuzzy logic to clustering.

He was an Assistant Professor at School of Economics, University “Dzemal Bijedic” and University of Sarajevo, from 1990 to 1993. From 1993 to 1994, he was Visiting Research, Tilburg University, the Netherlands. From 1994 to 1998, he was an Instructor of mathematics at the University of Georgia, Athens. In 1998, he was an Assistant Professor of mathematics at “James Madison” University in Harrisonburg, Harrisonburg, VA. From 1999, he works at School of Medicine, Emory University. He has more than 200 conference and journal publications, almost all of them in field of medical outcomes and its consequences.

Maysam Ghovanloo (S’00–M’04–SM’10) was born in 1973 in Tehran, Iran. He received the B.S. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 1994, the M.S. degree in biomedical engineering from the Amirkabir University of Technology,
Tehran, Iran, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, in 2003 and 2004, respectively.

He was an Assistant Professor in the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, from 2004 to 2007. In June 2007, he joined the faculty of Georgia Institute of Technology, Atlanta, where he is currently an Associate Professor and the Founding Director of the GT-Bionics Laboratory in the School of Electrical and Computer Engineering. He has authored or coauthored more than 90 peer-reviewed conference and journal publications.

Dr. Ghovanloo is an Associate Editor of the IEEE Transactions on Circuits and Systems, Part-II, IEEE Transactions on Biomedical Circuits and Systems, and a member of the Imagers, MEMS, Medical, and Displays (IMMD) subcommittee at the International Solid-State Circuits Conference (ISSCC). He a recipient of a CAREER Award from the National Science Foundation in 2010. He has also received awards in the 40th and 41st Design Automation Conference/ISSCC Student Design Contest in 2003 and 2004, respectively. He has organized several special sessions and was a member of Technical Review Committees for major conferences in the areas of circuits, systems, sensors, and biomedical engineering. He is a member of the Tau Beta Pi, AAAS, Sigma Xi, and a Senior Member of the IEEE Solid-State Circuits Society, IEEE Circuits and Systems Society, and IEEE Engineering in Medicine and Biology Society.
Fig. 1.
(a) TDS prototype comprised of a headgear with an array of lateral magnetic sensors, a wireless control unit, a USB receiver dongle, and a small magnetic tracer embedded in a titanium tongue stud with M&M-shaped upper ball. (b) Experimental setup with the subject sitting 1 m away from a 22” LCD monitor, performing the multidirectional tapping task.
Fig. 2.
GUI screen for (a) horizontal and vertical unidirectional tappings, (b) center-out tapping with all 48 possible target conditions on the right panel, and (c) multidirectional tapping with all 45 possible target conditions and their sequential order of tapping on the right panel.
Fig. 3.
(a) Recommended tongue positions for six TDS tongue commands plus the tongue resting position, which is considered neutral. (b) Designated keys on the keypad to resemble the TDS commands' positions.
Fig. 4. Unidirectional tapping tasks: (a) throughput, (b) error rate, and (c) task completion time of horizontal tapping and (d) throughput, (e) error rate, and (f) task completion time of vertical tapping.
Fig. 5.
Center-out tapping task: (a) throughput, (b) error rate, (c) task completion time, (d) path efficiency, (e) deviation from optimum number of movements, and (f) reaction time.
Fig. 6. Multidirectional tapping task: (a) throughput, (b) error rate, (c) task completion time, (d) path efficiency, and (e) deviation from optimum number of movement.
Fig. 7.
Sample cursor traces of one subject in the center-out tapping task with (a) TDS in the first session (TP = 0.91 bits/s, error rate = 12.5%, TCT = 7.4 min, PE = 72.30%, DONM = 1.59, and RT = 0.64 s); (b) TDS in the fifth session (TP = 1.87 bits/s, error rate = 15.97%, TCT = 5.3 min, PE = 84.40%, DONM = 0.19, and RT = 0.61 s); (c) keypad in the fifth session (TP = 2.10 bits/s, error rate = 7.64%, TCT = 5.2 min, PE = 85.70%, DONM = 0.01, and RT = 0.53 s); and (d) mouse in the first session (TP = 3.88 bits/s, error rate = 1.39%, TCT = 2.1 min, PE = 65.70%, DONM = 0.09, and RT = 0.10 s).
Fig. 8.
Comparison between the TDS and the keypad in the fifth session in terms of (a) effective distance, (b) effective width, (c) movement time in the center-out task, and (d) movement time in the horizontal tapping task.
Fig. 9.
Comparison between the TDS and the keypad in the fifth session of the center-out task in terms of movement time MT versus effective index of difficulty $ID_e$ and their regression models based on Fitts' law.
### TABLE I
Indices of Difficulty in Unidirectional (U) and Center-Out (C) Tasks

<table>
<thead>
<tr>
<th>Pixels</th>
<th>D1U</th>
<th>D2U</th>
<th>D3U</th>
<th>D1C</th>
<th>D2C</th>
<th>D3C</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>244</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>488</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1</td>
<td>30</td>
<td>2.34</td>
<td>3.19</td>
<td>4.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W2</td>
<td>61</td>
<td>1.59</td>
<td>2.32</td>
<td>3.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W3</td>
<td>122</td>
<td>1.00</td>
<td>1.59</td>
<td>2.32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE II
Indices of Difficulty in Multidirectional (M) Task

<table>
<thead>
<tr>
<th>Pixels</th>
<th>D1M</th>
<th>D2M</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1</td>
<td>57</td>
<td>2.67</td>
</tr>
<tr>
<td>W2</td>
<td>76</td>
<td>3.00</td>
</tr>
</tbody>
</table>
TABLE III

Summary of the Result for All Tapping Tasks

<table>
<thead>
<tr>
<th>Task</th>
<th>performance Measures</th>
<th>Mouse</th>
<th>TDS in 1st session</th>
<th>TDS plateau level (session)</th>
<th>TDS in 5th session</th>
<th>TDS 1st and 5th session significance (p)</th>
<th>KP in 5th session</th>
<th>TDS-KP 5th session significance (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Horizontal tapping</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP (b/s)</td>
<td></td>
<td>4.2 ± 1.3</td>
<td>2.1 ± 0.3</td>
<td>2.5 ± 0.2 (3)</td>
<td>2.5 ± 0.4</td>
<td>0.011</td>
<td>3.3 ± 0.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ER%</td>
<td></td>
<td>15.4 ± 21.6</td>
<td>24.7 ± 8.6</td>
<td>19.5 ± 5.8 (3)</td>
<td>15.8 ± 7.4</td>
<td>0.004</td>
<td>5.4 ± 3.0</td>
<td>0.002</td>
</tr>
<tr>
<td>TCT (min)</td>
<td></td>
<td>1.5 ± 0.4</td>
<td>3.9 ± 0.5</td>
<td>3.9 ± 0.5 (1)</td>
<td>3.7 ± 0.3</td>
<td>0.109*</td>
<td>3.2 ± 0.1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Vertical tapping</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP (b/s)</td>
<td></td>
<td>4.7 ± 1.2</td>
<td>2.2 ± 0.4</td>
<td>___ **</td>
<td>2.7 ± 0.2</td>
<td>0.002</td>
<td>3.2 ± 0.3</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ER%</td>
<td></td>
<td>11.4 ± 17.9</td>
<td>23.0 ± 8.8</td>
<td>___ **</td>
<td>13.3 ± 4.8</td>
<td>0.004</td>
<td>4.5 ± 3.7</td>
<td>0.014</td>
</tr>
<tr>
<td>TCT (min)</td>
<td></td>
<td>1.7 ± 0.5</td>
<td>3.8 ± 0.5</td>
<td>3.8 ± 0.5 (1)</td>
<td>3.7 ± 0.2</td>
<td>0.423*</td>
<td>3.2 ± 0.1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Center-Out tapping</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP (b/s)</td>
<td></td>
<td>3.7 ± 0.6</td>
<td>0.7 ± 0.2</td>
<td>1.4 ± 0.2 (3)</td>
<td>1.5 ± 0.2</td>
<td>&lt; 0.001</td>
<td>2.0 ± 0.2</td>
<td>0.001</td>
</tr>
<tr>
<td>ER%</td>
<td></td>
<td>5.1 ± 5.5</td>
<td>33.5 ± 9.1</td>
<td>18.4 ± 7.4 (3)</td>
<td>14.8 ± 7.0</td>
<td>0.008</td>
<td>4.9 ± 2.7</td>
<td>0.003</td>
</tr>
<tr>
<td>TCT</td>
<td></td>
<td>1.7 ± 0.5</td>
<td>7.3 ± 0.6</td>
<td>6.6 ± 0.6 (2)</td>
<td>6.2 ± 0.5</td>
<td>0.001</td>
<td>5.7 ± 0.3</td>
<td>0.001</td>
</tr>
<tr>
<td>PE%</td>
<td></td>
<td>66.1 ± 7.5</td>
<td>73.3 ± 4.3</td>
<td>76.6 ± 5.2 (2)</td>
<td>79.0 ± 3.1</td>
<td>0.021</td>
<td>84.9 ± 0.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>DONM</td>
<td></td>
<td>N/A</td>
<td>1.1 ± 0.5</td>
<td>1.1 ± 0.5 (1)</td>
<td>0.6 ± 0.3</td>
<td>0.082*</td>
<td>0.1 ± 0.1</td>
<td>0.001</td>
</tr>
<tr>
<td>RT (sec)</td>
<td></td>
<td>0.1 ± 0.04</td>
<td>0.8 ± 0.1</td>
<td>0.7 ± 0.1 (2)</td>
<td>0.7 ± 0.1</td>
<td>0.027</td>
<td>0.6 ± 0.1</td>
<td>0.035</td>
</tr>
<tr>
<td><strong>Multi-Directional tapping</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP (b/s)</td>
<td></td>
<td>4.5 ± 0.7</td>
<td>0.4 ± 0.1</td>
<td>0.9 ± 0.3 (4)</td>
<td>1.0 ± 0.2</td>
<td>&lt; 0.001</td>
<td>1.2 ± 0.2</td>
<td>0.009</td>
</tr>
<tr>
<td>ER%</td>
<td></td>
<td>3.5 ± 3.3</td>
<td>22.7 ± 8.2</td>
<td>___ **</td>
<td>6.9 ± 5.2</td>
<td>&lt; 0.001</td>
<td>2.0 ± 2.1</td>
<td>0.007</td>
</tr>
<tr>
<td>TCT (min)</td>
<td></td>
<td>1.8 ± 0.4</td>
<td>11.0 ± 1.3</td>
<td>9.3 ± 1.1 (2)</td>
<td>9.0 ± 1.2</td>
<td>0.001</td>
<td>8.5 ± 0.6</td>
<td>0.091*</td>
</tr>
<tr>
<td>PE%</td>
<td></td>
<td>91.5 ± 3.0</td>
<td>69.3 ± 3.3</td>
<td>73.9 ± 3.1 (2)</td>
<td>74.8 ± 3.7</td>
<td>0.001</td>
<td>78.9 ± 0.7</td>
<td>0.004</td>
</tr>
<tr>
<td>DONM</td>
<td></td>
<td>N/A</td>
<td>1.5 ± 0.5</td>
<td>0.8 ± 0.6 (2)</td>
<td>0.8 ± 0.6</td>
<td>&lt; 0.001</td>
<td>0.2 ± 0.1</td>
<td>0.005</td>
</tr>
</tbody>
</table>

*Not significantly different.

**No plateau was observed.