Published in final edited form as:

IEEE Trans Neural Syst Rehabil Eng. 2013 November ; 21(6): 979–991. doi:10.1109/TNSRE.2013.2248748.

A Dual-Mode Human Computer Interface Combining Speech and Tongue Motion for People with Severe Disabilities

Xueliang Huo [Member, IEEE],

GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308 USA. He is currently with the Interactive Entertainment Business of Microsoft, Redmond, WA 98052 USA

Hangue Park [Student Member, IEEE],

GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308 USA

Jeonghee Kim [Student Member, IEEE], and

GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308 USA

Maysam Ghovanloo [Senior Member, IEEE]

GT-Bionics Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308 USA

Maysam Ghovanloo: mgh@gatech.edu

Abstract

We are presenting a new wireless and wearable human computer interface called the dual-mode Tongue Drive System (dTDS), which is designed to allow people with severe disabilities to use computers more effectively with increased speed, flexibility, usability, and independence through their tongue motion and speech. The dTDS detects users' tongue motion using a magnetic tracer and an array of magnetic sensors embedded in a compact and ergonomic wireless headset. It also captures the users' voice wirelessly using a small microphone embedded in the same headset. Preliminary evaluation results based on 14 able-bodied subjects and three individuals with high level spinal cord injuries at level C3–C5 indicated that the dTDS headset, combined with a commercially available speech recognition (SR) software, can provide end users with significantly higher performance than either unimodal forms based on the tongue motion or speech alone, particularly in completing tasks that require both pointing and text entry.

Index Terms

Assistive technologies; human computer interaction; speech recognition; spinal cord injury; tongue motion; wireless

I. Introduction

Individuals with severe disabilities, such as those paralyzed as a result of spinal cord injuries (SCI) at levels C4 and above, stroke, amyotrophic lateral sclerosis (ALS), or traumatic brain injuries (TBI), heavily rely on assistive technologies (AT) to carry out various tasks in their everyday lives. Among ATs, those providing alternative control for computer access and wheeled mobility are considered the most important for today's active lifestyle since they can improve the users' quality of life (QoL) by easing two major limitations: effective communication and independent mobility [1], [2].

Computers and internet are regarded as great equalizers that allow all individuals to have similar vocational and recreational opportunities. It is generally accepted that once an individual with a disability is "enabled" to move around and effectively access computers or smartphones, he/she can virtually do most of the things that able-bodied individuals with educational, administrative, or scholarly careers do on a daily basis [3], [4]. This has resulted in a considerable amount ongoing research towards developing new ATs that can potentially take advantage of any remaining abilities of these individuals, such as head motion [5], eye movements [6], muscle contractions [7], and even brain signals [8], [9], to provide this population with alternative means to interact with computers and electronic gadgets.

The tongue, as one of the most flexible and capable parts of human body [10], has been considered as a suitable candidate for sophisticated motor control tasks by various researchers in the fields of AT and human computer interaction (HCI) [11], [12]. By taking advantage of the rich capabilities of tongue, we have already developed a wireless and wearable tongue-operated human computer interface, called the Tongue Drive System (TDS), which can enable individuals with severe physical disabilities to control their environments, access computers, and drive powered wheelchairs through their volitional tongue movements [13]–[16]. TDS detects the tongue motion by measuring the magnetic field variation generated by a magnetic tracer attached to the tongue using an array of magnetic sensors mounted on a wireless headset [13].

The TDS performance as a pointing device for computer access has been quantitatively and comparatively (versus keypad) evaluated on able-bodied subjects and reported in [17] and [18]. For example, the throughput of TDS in conducting ISO9241-9 center-out tapping tasks is 1.5 bits/s compared with 2.0 bits/s of keypad [17]. In our first clinical trial, we also demonstrated that an important group of the TDS potential end users, i.e., individuals with high level SCI (C2–C5), can use this system to substitute the mouse function in moving the mouse cursor to complete computer tasks or replace a manual joystick to navigate a powered wheelchair through an obstacle course [15]. These experiments have proven that the TDS can independently provide its users with a mean to access computers, navigate a wheelchair, and control their environments.

It has been understood, however, that an interface that is designed around only one input modality may not be fast and flexible enough to meet the diverse needs of the end users in today's hectic and demanding lifestyles [19]. Most existing interface devices, including our

TDS, operate well for a narrow set of specific tasks, under a specific set of environmental conditions, for users with a specific set of remaining abilities. Due to the wide variety of tasks in daily life, various types and levels of disabilities, the multitude of environmental conditions, and diversity of user goals and preferences, ATs that work perfectly well for one set of tasks, users, and environments, might show a poor performance in other tasks or environments by the same users or even completely lose their functionality when used for other applications by other users. In addition to the environmental and operating conditions, the performance of the single-mode ATs can be further degraded by the users' condition, such as fatigue, spasms, weakness, accent, etc.

A multimodal interface that expands the user access beyond one input channel, on the other hand, can potentially improve the speed of access by increasing the information transfer bandwidth between users and computers [20], [21]. A clear proof of this fact is the use of both mouse/touchpad and keyboard by the majority of able-bodied users on their desktop or laptop machines. In addition, multimodal human computer interfaces increase the number of alternatives available to users to accomplish a certain task, giving users the ability to switch among different input modalities, based on their convenience, familiarity, and environmental conditions [22]. Multimodal interfaces can also provide their users with more options to cope with fatigue. This is an important factor that improves the acceptability of ATs and can result in greater user satisfaction and technology adoption.

The TDS in its current form has been mainly designed to substitute mouse cursor movements in cardinal directions plus clicking functions by offering users six simultaneously accessible commands associated with particular user-defined positions in the mouth, which are activated when they are reached by their tongues. Even though TDS can provide full typing capability when it is used with an on-screen keyboard, its relatively small number of discrete commands, when compared to a full keyboard, and fast response time (< 0.5 s) makes it more suitable for mouse cursor control as opposed to typing [13].

On the other hand, speech recognition (SR) technology can offer a nearly unlimited number of available commands and has been regarded as one of the most efficient ways for text entry [23]. Individuals with severe disabilities can benefit from this technology as long as their vocal abilities are intact. Most SR software packages also allow their user to control the mouse cursor using a set of predefined voice commands. However, prior studies have shown that the task completion times for cursor control tasks are unacceptably long due to the inherent delay of SR [24]. Moreover, the ambient acoustic noise can significantly degrade the quality of sound acquired by the microphone and affect the accuracy of the SR interface. As a result, a system that relies on the speech input alone might show poor performance in translating users' rapid verbal commands or becomes completely irresponsive in noisy and outdoors environments.

The main objective of the work presented here is to combine the TDS and SR technologies to create a unified and highly integrated interface, called dual-mode Tongue Drive System (dTDS), which can take advantage of the strength of each modality to provide people with severe disabilities with a more efficacious, flexible, and reliable computer access tool that can be used in a wider variety of personal and environmental conditions. We expect the

dTDS to help its users in the following ways: 1) Increasing the speed of access by using each modality for its optimal target tasks and functions; 2) Allowing users to select either technology depending on the personal and environmental conditions, such as weakness, fatigue, acoustic noise, and privacy [20]; 3) Provide users with a higher level of independence by eliminating the need for switching from one AT to another, which often requires receiving assistance from a caregiver.

II. Dual-Mode Tongue Drive System (dTDS)

The dTDS, the block diagram of which is shown in Fig. 1, operates based on the information collected from two independent input channels: free voluntary tongue motion and speech. The two input channels are processed independently, while being simultaneously accessible to the users.

The primary dTDS modality involves tracking tongue motion in the 3-D oral space using a small magnetic tracer attached to the tongue via adhesives, piercing, or implantation and an array of magnetic sensors, similar to the original TDS. Detailed operational principle of the TDS has been described in [13]–[15].

The secondary dTDS input modality is based on the user's speech, captured using a microphone, conditioned, digitized, and wirelessly transmitted to the smartphone/PC along with the magnetic sensor data. Both TDS and SR modalities are simultaneously accessible to the dTDS users, particularly for mouse navigation and typing, respectively, and they have the flexibility to choose their desired input mode for any specific task without external assistance. The tongue-based primary modality is always active and regarded as the default input modality. The tongue commands, however, can be used to enable/disable the speech-based secondary modality via the dTDS graphical user interface (GUI) to reduce the system power consumption and extend battery lifetime.

III. dTDS Prototype

The latest dTDS prototype, built on a customized wireless headset, is an enhanced version of original TDS with the necessary hardware for a two-way wireless audio link to acquire and transmit users' vocal commands, while providing them with auditory feedback through an earphone. Fig. 2 shows the main components of the dTDS prototype, including: 1) A small permanent magnetic tracer attached to the tongue using tissue adhesives or embedded in a titanium tongue stud; 2) A custom-designed wireless headset, fabricated through 3-D rapid prototyping, which mechanically supports an array of four three-axial magnetic sensors and a microphone plus their interfacing circuitry to measure magnetic field and acoustic signals. A control unit combines and packetizes the acquired raw data before wireless transmission; 3) A wireless transceiver that receives the data packets from the headset and delivers them to the PC or smartphone; and 4) A GUI running on the PC or smartphone that includes high throughput data communication drivers and a sensor signal processing (SSP) algorithm that classifies the magnetic sensor signals and delivers the acoustic data to a commercial SR software.

A. Permanent Magnetic Tracer

A small ($\varnothing 3 \text{ mm} \times 1.6 \text{ mm}$) disc-shaped rare earth magnet (K&J Magnetics, Jamison, PA) with high residual magnetic strength ($B_r = 14,500 \text{ Gauss}$) was used as the tracer. A small tracer is desired to minimize any risk of discomfort and potential impact on the user's speech, which is important in achieving high accuracy with commercial SR software. The high B_r resulted in maintaining the signal-to-noise (SNR) despite shrinking the magnetic tracer.

B. Wireless Headset

A customized wireless headset was designed to combine aesthetics with user comfort, mechanical strength, and stable positioning of the sensors. The headset was also designed to offer flexibility and adjustability to adapt to the user's head anatomy, while enabling proper positioning of the magnetic sensors and the microphone near the user's cheeks [25].

The headset, block diagram of which is shown in Fig. 3, has a pair of adjustable sensor poles, each of which holds a pair of three-axial magneto-impedance (MI) sensors (AMI306, Aichi Steel, Japan) near the subjects' cheeks, symmetrical to the sagittal plane. The MI sensors make use of the giant magneto-impedance effect, in which the impedance of sensing element, which is a magnetic amorphous metal wire, changes in response to the magnetic field strength [26]. Each sensor has three orthogonal sensing elements to measure the magnetic field vector in 3-D. It also includes a magnetic signal readout circuitry, a programmable gain amplifier (PGA), a 12-bit analog-to-digital converter (ADC), and an inter-integrated circuit (I2C) serial interface.

A low-power microcontroller (MCU) with a built-in 2.4 GHz RF transceiver (CC2510, Texas Instruments, Dallas, TX) communicates with each sensor through the I2C interface to acquire samples at 50 Hz, while turning on only one sensor at a time to save power. When all four sensors are sampled, the results are packed into one magnetic data frame to be ready for RF transmission.

The acoustic signal acquisition is managed by an audio codec (TLV320-AIC3204, TI, Dallas, TX) and delivered through the built-in inter-IC sound (I2S) interface of the CC2510 MCU [27]. A miniaturized SiSonic MEMS microphone (Knowles, Itasca, IL) was placed near the tip of the right sensor board, as shown in Fig. 2, to capture the acoustic signal. Digitized audio samples are compressed to an 8 bit format using the CC2510 built-in μ -Law compression hardware to save the RF bandwidth. Once a complete audio data frame consisting of 54 samples has been acquired in 6.75 ms, the MCU assembles an RF packet containing one audio and one magnetic data frame and transmits it wirelessly [27].

After sending each RF packet, the MCU expects to receive a back telemetry packet including one data frame and one optional audio frame, which depends on whether the uplink audio channel from the transceiver to the headset has been activated or not. The data frame contains control commands from the PC/smartphone to switch on/off the speech modality. The audio frame in the back telemetry packet contains digitized sound signals from the PC/smartphone. The MCU extracts the audio samples from the back telemetry packet and sends them to the playback DAC of the audio codec through I2S interface to

generate audible analog audio signals if the user attaches an earphone to the headset audio jack. The CC2510 MCU can handle an RF data rate of 500 kb/s, which is sufficient for bidirectional data and audio transmission.

A simple but effective wireless handshaking has been implemented between the headset and the wireless transceiver to establish a dedicated wireless connection between the two devices without interference by other nearby dTDS headsets. When the dTDS headset is turned on, it enters an initialization mode by default and broadcasts a handshaking request packet containing specific header and its unique network ID using a basic frequency channel (2.45 GHz) at 1-s time intervals for 1 min. If the headset receives a handshaking response packet back from a nearby USB transceiver within the initialization period, it will update its frequency channel, standby threshold, and other operating parameters which are included in the response packet. Then it sends an acknowledgement packet back to the transceiver to complete the handshaking. The headset then switches to normal operating mode using the received parameters. Otherwise, in the absence of a handshaking, the headset will enter the standby mode by blinking a red LED to indicate that the initialization has failed and the power cycle should be repeated.

Power management circuitry includes a miniaturized 130 mAh Lithium-Polymer battery, a voltage regulator, a low voltage detector, and a battery charger. dTDS consumes either 6 or 35 mA from a 3-V supply depending on whether the bidirectional audio channel is off or on. This would allow the system to be used continuously for 20 or 4 hours in the uni-modal TDS or dTDS modes, respectively. Table I summarizes some of the key features of the dTDS prototype.

C. Wireless USB Transceiver

Fig. 2 shows a prototype of the transceiver equipped with a USB port and two audio jacks to interface the magnetic sensor data and acoustic signals with the PC, respectively. The transceiver has two operating modes: handshaking and normal. In the handshaking mode, the transceiver first listens to any incoming handshaking request packets from dTDS headsets within range (~10 m). If the transceiver receives a handshaking request packet with an appropriate header and a valid network ID, it will scan through all available frequency channels and chooses the least crowded one as the communication channel for that specific headset. The transceiver then switches to transmit mode and sends a handshaking response packet to the headset, before switching back to receiver mode and waiting for the confirmation of the acknowledgement packet. If an acknowledge is received within 5 s, the transceiver will update its frequency channel to the same frequency as the dTDS headset channel and enters the normal operating mode to receive regular magnetic/audio data packets. Otherwise, the transceiver will notify the PC/smartphone that the handshaking has failed.

In the normal mode, the transceiver works like a bidirectional wireless gateway to exchange data and audio samples between the dTDS headset and the PC/smartphone. The magnetic data within the headset packets are extracted and delivered to the PC/smartphone through the USB port. The audio data, however, is streamed into a playback audio codec via its I2S interface and converted to an analog audio signal, which is then delivered to the microphone

input of the PC/smartphone through a 3.5 mm audio jack (see Fig. 2). The transceiver can also receive analog audio output from the PC/smartphone headphone jack and digitize it using the same audio codec and I2S interface. These audio samples are compressed using the CC2510 built-in μ -law compression hardware and packaged in an audio frame. Transceiver also receives data packets from the computer, which contain the dTDS operating parameters, used to program the dTDS headset on the fly. The data packet is combined with the audio frame to form a back telemetry RF packet which is then wirelessly sent back to the headset.

D. Graphical User Interface (GUI)

The current dTDS GUI has been developed in the LabVIEW environment and used mainly for testing and demonstration purposes. The SSP engine, which needs to process the magnetic sensor data in real time, has been implemented in C. Generally, there is no need to present the dTDS users with its GUI. As long as the SSP engine is running in the background, the dTDS can be used to directly substitute the mouse and keyboard functions in the Windows operating system to provide the user with access to all the applications or software on the PC. GUIs for major smartphone operating systems are currently under development [16].

For detecting the tongue commands, the SSP algorithm uses the K-nearest-neighbors (KNN) classifier to identify the incoming magnetic sensor samples based on their features, which are extracted through Principal Components Analysis (PCA) from the data that is collected during a training step prior to using the system [13]. The current dTDS prototype supports six individual tongue commands, including four directional commands (UP, DOWN, LEFT, and RIGHT) and two selection commands (LEFT-SELECT and RIGHT-SELECT) that are simultaneously available to the user plus a neutral command defined as the tongue resting position.

Any piece of commercially available or customized SR software that works with a regular microphone can be used with the dTDS, because the audio signals are directly applied to the microphone input of the computer or smartphone. Dragon Naturally Speaking (Nuance, Burlington, MA) was our choice in this prototype because of its lower cost, wide adoption by the disability community, and supporting various platforms (Windows, Mac, iPhone, etc.).

IV. Performance Evaluation

To evaluate the dTDS performance for computer access, we recruited both able-bodied subjects and individuals with high level SCIs to participate in a trial, which required them to perform a comprehensive set of tasks that involved both mouse cursor navigation and typing using either one or both modalities of the dTDS. We have included both able-bodied subjects and people with high-level SCI in this evaluation for two main reasons: 1) to form a good understanding about the dTDS's performance in comparison with the gold standard combination of mouse-keyboard by having able-bodied subjects perform the same tasks using both solutions; 2) to use the able-bodied subjects' results to generate baseline values to gauge the results of people with severe disabilities.

A. Subject Population

Able-bodied subjects: 14 able-bodied subjects (age: 21–30 years old, nine males and five females) were recruited from the Georgia Institute of Technology graduate and undergraduate student population. Seven subjects had prior experiences with the TDS. However, none of them were regular TDS users. They were included in this study to explore the effect of prior TDS experience on the dTDS performance. The other seven subjects were naive with respect to TDS. Two subjects had several hours (<10) of prior experience with the Dragon, and others had never used the Dragon SR software before. There were an equal number of native (7) and non-native (7) English speakers in the able-bodied subject pool.

SCI subjects: Three subjects with high level SCIs (C3–C5) participated in this study. They were all male and aged from 37–53 years old. All three subjects had received magnetic tongue piercings and were familiar with the TDS as a result of participating in our prior studies [15], [28]. Two SCI subjects were active computer users via Head-tracker, Dragon, and mouth stick. The other subject was not a regular computer user and completely new to Dragon. The necessary approval was obtained from the Institutional Review Board (IRB) of the Georgia Institute of Technology, and informed consent forms were collected from all subjects prior to the experiments.

B. Experimental Design

The purpose of the experiment was to compare the subjects' performance with the dTDS in completing comprehensive computer access tasks, involving both mouse cursor navigation and typing, with the unimodal counterparts (TDS and Dragon), and also to reveal important learning effects in using the dTDS on a short term basis. A within-subject model was designed with each subject repeating the same tasks using three devices: TDS, Dragon, and dTDS.

For able-bodied subjects, the experiment was performed in the GT-Bionics lab and divided into two sessions: instructional session and experimental session with a maximum of one week gap in between. For the SCI subjects, the experiment was conducted in their home with at least one of the subjects' care-givers or family members present. It included one instructional and six experimental sessions, with a maximum of four days gap in between two consecutive sessions. The general design of the experiment was the same for able-bodied and SCI subjects, and included following tasks: 1) a text transcription task, which required subjects to transcribe several short paragraphs from a hard copy onto a word document using two different microphones with the SR software; 2) a modified maze navigation task, in which subjects were asked to navigate the mouse cursor through an on-screen maze, click on designated areas, and type; 3) a modified ISO9241-9 center-out tapping task, in which the target clicking was randomly interleaved with typing tasks [29]. In the instructional session, subjects learned how to use the Dragon software and dTDS and familiarized themselves with the contents and flow of the experiment by going through all the experimental tasks only once. During the experimental session, subjects performed and repeated all tasks except the text transcription for four times using three different input devices: TDS, Dragon, and dTDS. Each session took ~3 hours.

C. Experimental Setup

During the experiment, shown in Fig. 4, subjects had a small permanent magnetic tracer attached to their tongue either temporarily using tissue adhesive (able-bodied) or semi-permanently via tongue piercing (SCI), worn a dTDS headset, and sat ~1 m away from a 22" monitor with 1280 × 800 resolution. To define dTDS tongue commands, they went through calibration, command identification, and training steps as explained in [13] and [14]. In the instructional session, subjects started with two-command training and practicing, then the number of commands was increased to four and then to six in order to help them gradually learn how to use the TDS from easy to advanced modes. In the experimental session, subjects started the session by defining and training the TDS with six tongue commands.

Detailed instructions on how to use the TDS commands to navigate the mouse cursor, issue left/right clicks, and type or erase letters were given during the instructional session and repeated during experimental sessions if needed. Briefly, to move the mouse cursor, subjects should issue a directional TDS command by touching a specific tooth that they have trained the SSP for that direction with the tip of their tongue. The speed of the cursor ramps up and reaches a predefined maximum value as the subjects hold their tongue tip on that tooth. The cursor stops immediately if the tongue is returned back to its resting position. Left/right clicks are issued by touching the left/right cheeks with the tip of the tongue for 0.5 s.

To type with TDS, subjects move the cursor on an on-screen keyboard in four directions to highlight letters or symbols. If subjects keep issuing a directional tongue command, the highlight on the keyboard moves in that direction in 0.5-s intervals. Touching the left cheek with the tongue tip for 0.5s will type the highlighted letter, while touching the right cheek for 0.5s will delete the last entered character.

Subjects also received instructions on using the SR software if they had not used it before. For typing, subjects just simply uttered the words or sentences via the microphone in the same way that they had trained the software. For mouse navigation, a set of predefined verbal commands, such as “move mouse Left/Right/Up/Down,” “stop,” “move mouse slow,” “much faster,” and “left/right select” were used. Similarly, mouse clicks were issued by dictating those commands.

D. Tasks

Text Transcription—The purpose of this task was to evaluate the performance of the dTDS wireless microphone, incorporated in the dTDS headset, and to compare it with a commercial microphone that was designed specifically for the SR software. This experiment was only repeated once in the practicing session without any TDS function.

The subjects were provided with two ~120 word test paragraphs (The North Wind and the Sun and The Grandfather Passage [30]) and asked to read both loudly for three times for practice. The subjects trained the Dragon by reading ten short paragraphs, provided by the manufacturer, using a commercial microphone, NC-181VM (Andrea Electronics, New York, NY). Subjects were then asked to transcribe two test paragraphs from a hardcopy to a word document via dictation using the commercial microphone. Subjects were asked to dictate in their regular speed and not to correct errors. Then magnetic tracers were attached

to the subjects' tongues, and they repeated the Dragon training and text transcription using the wireless microphone function incorporated in the dTDS headset. Results were used to calculate the SR accuracy for that specific audio input device.

Maze Navigation—Subjects were instructed to move the cursor as quickly and as accurately as possible through an on-screen maze [see Fig. 5(a)] from the red START circle to the END circle, trying to keep the cursor within the blue track. For each trial, a new maze was randomly selected from five different patterns which had a similar level of difficulty with the same width and length of segments, but different orientations. At the beginning of each trial, the mouse cursor was automatically positioned in the center of the START circle and subjects were instructed to only start moving the cursor after red circle turned green. The experiment automatically terminated when subjects reached the END circle.

Subjects were also required to stop at the yellow and green checkpoints and issue Right or Left select commands, respectively, according to the letter inside each checkpoint. Upon selection, the checkpoint color changed to gray, indicating correct issuance of the designated command on that checkpoint. Out of 11 checkpoints, four were randomly associated with typing tasks. Once subjects clicked on such a checkpoint, a typing window, shown in Fig. 5(b), popped up and subjects were asked to type a short phrase shown on top of the window as quickly and accurately as possible. The typed characters appeared in a textbox below the target phrase. Subjects were asked to correct user-induced typing errors, such as subjects dictating wrong words, but ignore the system related errors, such as Dragon recognition errors. Four different types of representative phrases, including common phrases, websites, phone numbers, and spelled words, appeared one at a time in a randomized order in each round. Overall, the subjects had to complete a minimum of 12 cursor movements, 11 clicks (excluding those for typing with TDS), and on average, typed 36 characters in each round of the trial.

Modified Center-Out Tapping—Subjects were required to perform a modified ISO9241-9 center-out tapping task [29], in which clicking targets were randomly interleaved with typing tasks. ISO9241-9 is based on the well-known Fitts' law and widely adopted by the HCI community for quantitatively evaluating conventional non-keyboard input devices as well as ATs, such as eye trackers, head trackers, and voice activated software [31]–[34]. By adding typing to the center-out tapping task, we have embedded the two key features of web surfing in one task. Unlike web surfing, the resulting task is still quantifiable and closely resembles the real world application.

In the modified center-out tapping task, circular targets with three different diameters ($W = 30, 61, \text{ and } 122$ pixels) and three different distances from the center of the screen ($D = 61, 122, \text{ and } 244$ pixels) created a total of six $D - W$ pair. They appeared randomly on the screen one at a time along cardinal and ordinal directions (every 45° from the center), forming a total of 48 (6×8) targets, as shown in Fig. 5(c), similar to [18]. These 48 targets were evenly divided into three experiments each having 16 targets associated with two $D - W$ pairs in each round. Subjects were instructed to move the cursor from the center of the screen as soon as the targets appeared. They were also instructed to issue a left select as close to the center of the target as possible to confirm the selection. Four out of 16 targets in

each round were randomly associated with typing tasks, in which subjects had to complete a task similar to what was described in the maze navigation session.

Subjects were required to complete both maze navigation and center-out tapping tasks using the TDS alone, Dragon alone, and dTDS. During the experimental session, all tasks were repeated four times for each variation, one for practice followed by three testing rounds. When using the TDS, the microphone was turned off to deactivate Dragon. When using Dragon, the TDS functions were disabled. Typing and mouse navigation were both completed using the dTDS microphone along with Dragon software. In the dTDS mode, both the TDS and Dragon were active and subjects were required to use the tongue commands for mouse navigation and clicks, and verbal commands for typing. The order of using each device was randomized to minimize the learning effect. At the end, able-bodied subjects were also asked to perform both tasks with a combination of standard mouse and keyboard to generate a reference point.

Questionnaire—At the end of the last session, subjects answered a questionnaire rating different aspects of each device, such as speed, accuracy, being easy-to-learn, causing fatigue, etc. Able-bodied subjects were also asked about their preferred input modality to access a computer if they were not able to use a mouse or keyboard, while SCI subjects were asked about their preference between the dTDS and their current AT devices.

E. Performance Measures

The performance measure for text transcription experiment was recognition accuracy, which was defined as the percentage of correctly recognized words to the total number of words included in the two text reading paragraphs.

For maze navigation, the performance measures include total completion time, cursor navigation time, error-free (EF) typing rate, navigation error, and typing error. The first three measures indicate the speed of each input device in completing specific tasks. Total completion and cursor navigation time were calculated from the recorded time during the experiment. The EF typing rate was calculated from dividing the number of correctly typed characters during typing task by typing time. Navigation error is the summation of all the deviations of the cursor path from the edges of the blue track divided by 1000 as a measure of navigation accuracy. The typing error was calculated as the percentage of mistyped letters over the total number of letters to be typed during the typing tasks.

For the center-out tapping, in addition to total completion time, cursor navigation time, EF typing rate, and typing error, as those defined in the maze navigation, two more performance measures including throughput and error rate were considered to assess the speed and accuracy in pointing and selecting. Throughput (TP), measured in bits/s, is an indicator for the amount of information that users can deliver to a computer through the device under test. It is defined as the ratio between the effective Index of Difficulty, IDE, of targets with the same condition (i.e., same $D - W$ pair) to movement time (MT), the time it takes to reach them. Here we used the same definition of the IDE as in [18] and [35], while the MT was calculated from the time cursor starts moving to when the subject issues a select command. It neither includes the initiation delay time before the subject moves the cursor nor the

selection time. However, it does include the small break times when the subject halts the cursor before reaching the target. 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, ER reveals the subjects' accuracy in using the computer input device for pointing and selecting the targets [36].

V. RESULTS

All subjects successfully completed all sessions. Since the microphone performance was independent of the subjects abilities (none of the SCI subjects were on a ventilator), we have presented those results from all subjects combined. Other results, such as performance measures in comprehensive computer access tasks and questionnaires, are presented separately for the two subject groups. Performance measures for each device were calculated by first averaging within each individual subject across three testing rounds and then averaging across all subjects in each group. Unless stated otherwise, the data is presented as mean $\pm$ SD in the text and tables and mean $\pm$ SE in figures.

A. Performance of dTDS Microphone

Subjects achieved an overall $87.5\% \pm 8.8\%$ and $85.1\% \pm 8.8\%$ recognition accuracy using the commercial and dTDS microphones, respectively. These results were lower than the accuracy claimed by Nuance ($>95\%$), the manufacturer of Dragon Naturally Speaking software. This is possibly because of the seven non-native English speakers among subjects and unfamiliarity with the Dragon software. Although the order of the microphones was fixed, the impact of learning was expected to be minimal as the subjects had already practiced both paragraphs for a few times before the experiment. We ran a Two-way analysis of variation (ANOVA) on the recognition accuracy considering both device and English accent as factors. This analysis revealed that native speakers (10) performed significantly better than non-native speakers (7) ($93.8\% \pm 3.1\%$ and $91.7\% \pm 4.0\%$ versus $78.6\% \pm 5.9\%$ and $75.7\% \pm 5.1\%$ for commercial and dTDS microphones, respectively) with $p < 0.01$, while there was no significant difference between the performance of commercial and dTDS microphones. Dependency on the users' accents is one of the issues associated with unimodal voice control devices aside from their susceptibility to ambient acoustic noise.

B. Maze Navigation

Fig. 6 shows the averaged total completion time, cursor navigation time, and EF typing rate of the three input devices mentioned above, plus mouse/keyboard combination as reference, in completing the maze navigation experiment by able-bodied subjects. One-way ANOVA was conducted with device as a factor showing significant effect of device on all three measures. Pair-wised comparison with Bonferroni adjustment revealed that the total completion time of dTDS (92 ± 19 s) was significantly shorter than those of TDS (285 ± 69 s) and Dragon (202 ± 32 s) alone, with $p < 0.01$ in both cases. Same analysis also showed that there were significant differences between all performance measures of unimodal TDS and Dragon (all $p < 0.01$). TDS outperformed Dragon in terms of cursor navigation time (48 ± 18 s versus 146 ± 22 s), while Dragon was much faster in typing (39.6 ± 8.0 chars/min

versus 8.8 ± 2.3 chars/min). No significant difference was observed between the navigation time of the dTDS and TDS, and none between the EF typing rate of the dTDS and Dragon either. This shows that the TDS and SR (Dragon) can be used together in the dTDS without running the risk of degrading the user's performance due to their possible interference.

Table II summarizes the cursor navigation and typing errors using different devices in completing maze navigation tasks by able-bodied subjects. We performed one-way ANOVA on both accuracy measures with device as a factor. Results showed that there was no significant difference among the navigation errors of different devices except for mouse/keyboard. The typing error of the TDS is significantly lower than that of Dragon and dTDS ($p < 0.01$), while there was no difference between Dragon and dTDS.

For SCI subjects, in Fig. 7, the overall completion time for maze navigation tasks has reduced from 339 ± 81 s, 220 ± 5 s and 122 ± 46 s in the first session to 289 ± 66 s, 203 ± 31 s, and 90 ± 29 s in 6th session using TDS alone, Dragon alone, and dTDS, respectively, corresponding to 14.8%, 8.0%, and 26.5% improvements. More specifically, the navigation times are 61 ± 32 s versus 46 ± 15 s, 161 ± 11 s versus 141 ± 17 s, and 63 ± 35 s versus 37 ± 16 s, while the EF typing rates in the first versus sixth sessions are 8.0 ± 1.9 chars/min versus 9.4 ± 2.0 chars/min, 37.6 ± 1.4 chars/min versus 38.2 ± 6.7 chars/min, and 38.5 ± 6.2 chars/min versus 42.7 ± 9.9 chars/min, for TDS alone, Dragon alone, and dTDS, respectively. The most significant improvement was the navigation time of dTDS (41.3%), followed by the navigation time of TDS (25.9%) and EF typing rate of TDS (17.5%). At the end of the sixth session, the total completion time of using dTDS was only 31% and 44% of using TDS alone and Dragon alone, respectively.

C. Center-Out Tapping

Fig. 8 shows the averaged total completion time, cursor navigation time, and EF typing rate of all input devices in completing the center-out tapping task by able-bodied subjects. One-way ANOVA was conducted with device as a factor, showing a significant effect of device on all three measures. Pair-wised comparison with Bonferroni adjustment showed that the total completion time of dTDS (104 ± 16 s) was significantly shorter than those of TDS (282 ± 95 s) and Dragon (264 ± 30 s) alone, with the p -value < 0.01 in both cases. The same analysis also demonstrated that there were significant differences between the cursor navigation time ($p < 0.01$) and EF typing rate ($p < 0.01$) of unimodal TDS and Dragon, while there was no evidence of a difference between the total completion time of these two devices. Similar to the maze navigation, TDS's performance was better in navigation, while Dragon was much faster in typing. No significant difference was found between the navigation time of the dTDS and TDS, and none between EF typing rate (42.9 ± 8.7 chars/min versus 37.9 ± 7.0 chars/min) of the dTDS and Dragon.

Fig. 9 shows the average throughput and error rate in completing the center-out tapping task using different devices by able-bodied subjects. The mouse TP (4.44 ± 0.49 bits/s) is within the generally accepted range of 3.7–4.9 bits/s, which validates our methodology, GUI functionality, and data analysis [36]. One-way ANOVA considering a device as a factor showed a significant effect of the device on both measures. Pair-wised comparison with Bonferroni adjustment showed that the TP of both TDS (1.20 ± 0.38 bits/s) and dTDS (1.23

± 0.36 bits/s) are significantly higher than that of Dragon (0.38 ± 0.05 bits/s), both with $p < 0.01$, while there was no significant difference between the TP of dTDS and TDS. Similar results were observed for the ER. The ER of TDS ($20.1\% \pm 10.3\%$) and dTDS ($22.6\% \pm 10.4\%$) were significantly higher than that of Dragon ($5.1\% \pm 4.1\%$) with $p < 0.01$, but there was no significant difference between dTDS and TDS. The achieved TP and ER for dTDS and TDS were consistent with our previous findings [18].

Similarly, we performed one-way ANOVA on typing errors with a device as a factor. The overall typing error of the TDS ($0.8\% \pm 1.5\%$) was significantly lower than that of Dragon ($4.6\% \pm 3.3\%$) and dTDS ($5.4\% \pm 3.3\%$) with $p < 0.01$, while there was no significant difference between the typing error of Dragon and dTDS.

The overall completion time of center-out task by the SCI subjects did not change significantly when using Dragon and dTDS when comparing the results of the first and sixth sessions (269 ± 18 s versus 267 ± 43 s and 112 ± 30 versus 121 ± 42 s, respectively). As shown in Fig. 10, the performance improvement of using TDS alone is more pronounced (304 ± 81 s versus 272 ± 92 s in first and sixth session, respectively) compared to that of Dragon and dTDS. The EF typing rates in the first versus sixth sessions are 9.5 ± 2.5 chars/min versus 11.1 ± 2.6 chars/min, 37.5 ± 0.7 chars/min versus 42.4 ± 7.9 chars/min, and 43.1 ± 6.9 chars/min versus 39.3 ± 12.9 chars/min, while the navigation times are 62 ± 17 s versus 64 ± 26 s, 209 ± 16 s versus 212 ± 35 s, and 59 ± 19 s versus 58 ± 19 s, for TDS alone, Dragon alone, and dTDS, respectively. No improvement greater than 10% has been observed for these performance measures except for EF typing rate of the TDS (16.8%) and Dragon (13.1%). At the end of the sixth session, the total completion time of using dTDS was 44% and 45% of using TDS alone and Dragon alone, respectively.

D. User Perception

Able-bodied users' subjective ratings on different devices are summarized in Fig. 11. To minimize the bias, we only included the results of seven able-bodied subjects who were completely new to both TDS and Dragon. Responses were in a scale of 1 to 9, with a higher number representing a more positive perception about that specific topic. Statistical analysis (one-way ANOVA with pair-wised comparison) showed that dTDS received significantly higher ratings than the other two devices in terms of speed (both $p < 0.01$) and overall satisfaction (both $p < 0.01$). There was no significant difference among devices as far as the perceived accuracy was concerned. These results were consistent with the quantitative measurements mentioned above. All seven subjects reported that they prefer to use the dTDS over TDS or SR software if they were neither able to use mouse nor keyboard to access computers.

Fig. 12 shows the SCI subjects' rating on different devices at the end of the sixth session. dTDS received the highest scores in terms of accuracy, speed, ease-of-use, less thinking, and overall satisfaction. However, due to the limited number of subjects (three), the difference between dTDS and other two devices based on these subjective measures were not significant. The SCI subject who was a mouth stick user said that he preferred the dTDS over his current AT for computer access on a daily basis, while another SCI subject who was using a combination of head-tracer and Dragon reported that he liked his current ATs better

because of the proportional control with 360° directionality (similar to mouse) offered by the head-tracker. The third SCI subject who was not a regular computer user preferred to use dTDS over TDS or Dragon for computer access.

E. Effect of TDS Experience and English Accent

To examine the effect of different factors on the dTDS performance, we have categorized the able-bodied subjects based on their prior TDS experience (seven experienced and seven naive subjects) and accent (seven native and seven non-native English speakers). We have applied Two-way ANOVA on all the performance measures considering both factors. The results are summarized in Tables III and IV.

The effect of prior TDS experience is significant on the total completion time and the navigation time in both maze and center-out tasks. In addition, the average TP of TDS-experienced subjects was significantly higher than that of TDS-naive subjects in center-out tapping test (1.4 ± 0.3 versus 1.1 ± 0.3 bits/s). These results showed that the subjects with prior TDS experience have better control over mouse cursor when using the dTDS and can complete the tasks more efficiently compared with TDS novice. There was no significant effect of prior TDS experience on the navigation error in maze navigation and outside hitting error in the center-out tapping. Post-hoc power analysis showed that there was actually not enough power to detect that effect for these two performance measures due to the small number of subjects in each group. However, on average, the magnitude of both errors is considerably higher in less experienced subjects (4.0 ± 2.8 pixels versus 2.0 ± 2.0 pixels and $25.3\% \pm 12.6\%$ versus $19.9\% \pm 7.6\%$ for maze navigation and center-out error rates, respectively).

Large variations compared with the amplitude of the mean values reduced the effect size and resulted in low power in data analysis. Variations are expected to decrease with a higher number of subjects, which is expected to make the effect of TDS prior experience more pronounced.

Regarding the effect of accent, native English-speaking subjects typed significantly faster than non-native subjects in the modified center-out tapping task (47.4 ± 10.9 chars/min versus 38.4 ± 7.1 chars/min). Statistically, there was no significant effect of accent in the typing accuracy due to low power in the analysis as a result of the small number of subjects. However, the mean values of typing error for native English speakers ($3.4\% \pm 3.0\%$ and $5.1\% \pm 3.5\%$ in maze and center-out tapping, respectively) were much lower than non-native speakers ($4.7\% \pm 4.2\%$ and $7.7\% \pm 5.7\%$ in maze and center-out tapping, respectively). We observed an unexpected significant effect of accent on the navigation error in the maze navigation experiment. Basically, according to our results, non-native speakers can navigate significantly faster than native speakers. This was probably related to the fact that most of our TDS-experienced subjects (five out of seven) were non-native speakers.

VI. Discussion

The main purpose of this study was to compare the performance of TDS, SR software (Dragon) and dTDS in completing comprehensive computer access tasks and discover the

strengths and weaknesses of each modality. The results can be used to develop advanced ATs that will take advantage of the best modalities that are available to potential end users for achieving maximum performance.

We have developed the dTDS prototype to search for and overcome the key technical challenges in combining two input modalities, such as allocating enough bandwidth to transfer both audio and magnetic data, minimizing the electronic interference between the two modalities, and reducing the headset power consumption to extend the battery lifetime. It is desirable to use a realistic prototype as opposed to an emulated setup consisting of SR software, a separate microphone, and the unimodal TDS because the users' experience, which is among the most important factor affecting technology adaptation, would have been quite different.

In both popular computer access tasks, i.e., maze navigation and center-out tapping, which were modified to include typing, able-bodied subjects performed consistently faster in cursor navigation and typing using TDS and Dragon, respectively. Also, subjects obviously benefited from using the best of each modality in dTDS. This was evident from the lowest total completion time that was achieved when using the dTDS, which was 68% and 56% less time consuming than using the TDS or Dragon alone, respectively, in the maze navigation task, and 63% and 54% less time consuming in the center-out tapping task, respectively. Moreover, the total completion time of the dTDS was only ~3 times that of the mouse/keyboard combination. This is quite significant considering that all subjects had much less experience in using the dTDS compared to the mouse/keyboard combination, which they were using on a daily basis over many years. It is reasonable to expect the gap between the dTDS and mouse/keyboard to become smaller once subjects use the dTDS on a daily basis over an extended period of time.

Overall, using the dTDS could significantly increase the speed, but did not result in higher accuracy when considering both typing and navigation. This is mainly because of the large typing error rate associated with the SR modality, which seems to be more problematic for non-native English speakers. In real life, however, dTDS users do not have a mandate to only use the SR mode for typing. Occasionally, they can take advantage of the TDS modality for typing in order to achieve a higher level of accuracy, for instance when the SR software does not recognize their spoken input because of their accent or when there is a high level of acoustic noise in the environment.

We observed performance improvements in completing both maze navigation and center-out tasks by the SCI subjects using all three devices. However, the amount of improvement, i.e., reduction in total completion time, was not that significant compared to our previous findings in [18]. This is probably because our SCI subjects had already gained a significant amount of experience in using TDS for mouse navigation as a result of participating in our previous study. Also, two out of three subjects were using the Dragon software for daily computer access. Nevertheless, SCI subjects have achieved the most improvement in terms of dTDS navigation time and TDS navigation time in maze navigation, and TDS EF typing rate in both maze navigation and center-out tapping tasks, all of which are tasks that need to be completed by the TDS modality. In other words, even though the SCI subjects were

already familiar with TDS, their performance was still improving due to learning. It should be noted that the tasks in this study were different from those that the subjects had been exposed to in our previous studies. For example, in maze navigation, subjects were asked to click in designated areas and randomly type, which were not needed in our previous studies. In addition, typing with TDS was a completely new task that had not been tested before. As a result, the SCI subjects had to learn how to apply their TDS skills to accomplish these new tasks.

Similar to able-bodied subjects, the SCI subjects were able to accomplish the required tasks significantly faster using dTDS compared to its unimodal counterparts. At the end of the sixth session, the average performance of the SCI subjects was close to the significantly younger able-bodied subjects who had prior TDS experience (90s versus 83s for maze completion time, and 121s versus 96s for center-out completion time). This is evidence for the fact that the dTDS can be quickly learned and used for comprehensive computer access tasks regardless of the users' abilities, age, and education.

The typing task in both maze navigation and center-out typing experiments is not ideal and has its limitations. However, we have tried to make it as realistic as possible by requiring subjects to type four typical types of phrases, including common phrases (i.e., *How are you?*), websites (i.e., www.google.com), phone numbers (i.e., 404-368-7541), and spelled words (i.e., *EXKGSA*). These phrases are commonly used for online chatting, web surfing, phone dialing, and password setting in daily computer usage. We believe such an experiment can sufficiently represent the real-life typing tasks while considering the capabilities of both TDS and SR within the experiment time limits.

VII. Conclusion

We have developed a dTDS that allows people with severe disabilities to use computers by navigating a mouse cursor and typing via two modalities: voluntary tongue motion and speech, which are simultaneously available to them. The dTDS users can choose their preferred modality based on the nature of the tasks, operating environments, and their physical conditions. This can result in increased speed, flexibility, usability, and independence. Preliminary results from 14 able-bodied subjects and three subjects with high level SCI (C3–C5) showed that this new dual-mode human-computer interface can significantly improve the speed of completing complex computer access tasks, where both text entry and cursor navigation were necessary. It was also demonstrated that using the dTDS does not affect the user's performance with either one of the two modalities. Subjects with prior TDS experience were expectedly able to perform the tasks more efficiently than the novice subjects. We also observed the significant effect of accent on EF typing rate in the modified center-out tapping task. No significant effect of accent was discovered on the typing error due to the low power in ANOVA because of the limited number of subjects. All able-bodied subjects reported that they preferred using the dTDS over TDS or SR alone for computer access, while two out of three SCI subjects preferred the dTDS over their current ATs. We are planning to explore the addition of other input modalities, such as head control, to the current dTDS platform to further improve its performance and end user coverage. We

also intend to add proportional control capability to the TDS modality and make it more user-friendly in navigating the mouse cursor.

Acknowledgment

The authors would like to thank participants in their human subject trials, members of the GT-Bionics Lab for their assistance with conducting the trials, and Dr. Vidakovic for his constructive comments on statistical data analysis.

This work was supported in part by the National Institute of Biomedical Imaging and Bioengineering under Grant 1RC1EB010915 and by the National Science Foundation under Awards CBET-0828882 and IIS-0803184.

References

1. Carlson, D.; Ehrlich, N. Rep. U.S. Department of Education, Nat. Inst. Disability and Rehabilitation. Washington, D.C., USA: 2005. Assistive technology and information technology use and need by persons with disabilities in the United States.
2. National Institute of Neurological Disorders and Stroke (NINDS), NIH, Spinal cord injury: Hope through research [Online]. Available: http://www.ninds.nih.gov/disorders/sci/detail_sci.htm
3. Cook, AM.; Polgar, JM. Cook and Hussey's Assistive Technologies: Principles and Practice. 3rd ed.. St. Louis, MO, USA: Mosby; 2007.
4. Scherer, MJ. Living in the State of Stuck: How Assistive Technology Impacts the Lives of People With Disabilities. 4th ed.. Boston, MA, USA: Brookline; 2005.
5. Pereira C, Neto R, Reynaldo A, Luzo M, Oliveira R. Development and evaluation of a head-controlled human-computer interface with mouse-like functions for physically disabled users. Clin. Sci. 2009 Jul;64:975–981.
6. Barea R, Boquete L, Mazo M, Lopez E. System for assisted mobility using eye movements based on electrooculography. IEEE Trans. Rehab. Eng. 2002 Dec; 10(4):209–218.
7. Huang C, Chen C, Chung H. Application of facial electromyography in computer mouse access for people with disabilities. Disability Rehab. 2006 Feb;28:231–237.
8. Hochberg LR, et al. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature. 2006; 442:164–171. [PubMed: 16838014]
9. McFarland DJ, Krusienski DJ, Sarnacki WA, Wolpaw JR. Emulation of computer mouse control with a noninvasive brain—Computer interface. J. Neural Eng. 2008; 5(2):101–110. [PubMed: 18367779]
10. Kandel, ER.; Schwartz, JH.; Jessell, TM. Principles of Neural Science. 4th ed.. New York, NY, USA: McGraw-Hill; 2000.
11. Andreasen Struijk LNS. An inductive tongue computer interface for control of computers and assistive devices. IEEE Trans. Biomed. Eng. 2006 Dec; 53(12):2594–2597. [PubMed: 17152438]
12. Saponas S, Kelly D, Parviz BA, Tan DS. Optically sensing tongue gestures for computer input. Proc. ACM Symp. User Interface Software and Technology. 2009 Oct.;177–180.
13. Huo X, Wang J, Ghovanloo M. A magneto-inductive sensor based wireless tongue-computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. 2008 Nov; 16(5):497–504. [PubMed: 18990653]
14. Huo X, Wang J, Ghovanloo M. Using unconstrained tongue motion as an alternative control mechanism for wheeled mobility. IEEE Trans Biomed Eng. 2009 Jun; 56(6):1719–1726. [PubMed: 19362901]
15. Huo X, Ghovanloo M. Evaluation of a wireless wearable tongue—Computer interface by individuals with high-level spinal cord injuries. J. Neural Eng. 2010 Apr;7:497–504.
16. Kim J, Huo X, Minocha J, Holbrook J, Laumann A, Ghovanloo M. Evaluation of a smartphone platform as a wireless interface between Tongue Drive system and electric-powered wheelchairs. IEEE Trans. Biomedical Eng. 2012 Jun; 59(6):1787–1796.
17. Yousefi B, Huo X, Ghovanloo M. Using Fitts's law for evaluating tongue drive system as a pointing device for computer access. Proc. 32nd IEEE Eng. in Med. and Biol. Conf. 2011 Sep.: 4404–4406.

18. Yousefi B, Huo X, Ghovanloo M. Quantitative and comparative assessment of learning in a tongue-operated computer input device. *IEEE Trans. Info. Tech. BioMed.* 2011 Sep; 15(4):747–457.
19. Keates S, Robinson P. The use of gestures in multimodal input. *Proc. 3rd Intl. ACM Conf. on Assist. Tech.* 1998 Apr.:35–42.
20. Shein, F.; Brownlow, N.; Treviranus, J.; Pames, P. *Proc. Visions Conf.: Augmentative and Alternative Comm. in the Next Decade.* Wilmington, DE, USA: Univ. Delaware/Alfred I. duPont Inst; 1990. Climbing out of the rut: The future of interface technology.
21. Smith A, Dunaway J, Demasco P, Peichl D. Multimodal input for computer access and alternative communication. *Proc. 2nd Annu. ACM Conf. on Assist. Tech.* 1996 Apr.:80–85.
22. Baljko M. The contrastive evaluation of unimodal and multimodal interfaces for voice output communication aids. *Proc. 7th Intl. Conf. Multimodal Interfaces.* 2005 Oct.:301–308.
23. Rebman CM, Aiken MW, Cegielski CG. Speech recognition in the human-computer interface. *Inf. Manage.* 2003 Jul.40:509–519.
24. Dai L, Goldman R, Sears A, Lozier J. Speech-based cursor control: A study of grid-based solutions. *SIGACCESS Access. Comput.* 2003 Sep.:94–101.
25. Park H, Kim J, Huo X, Wang IO, Ghovanloo M. New ergonomic headset for tongue-drive system with wireless smartphone interface. *Proc. 33rd IEEE Eng. in Med. and Biol. Conf.* 2011 Sep.: 7344–7347.
26. Phan MH, Peng HX. Giant magnetoimpedance materials: Fundamentals and applications. *Progr. Materials Sci.* 2008; 53:323–420.
27. Huo X, Ghovanloo M. Using speech recognition to enhance the tongue drive system functionality in computer access. *Proc. 33rd IEEE Eng. in Med. and Biol. Conf.* 2011 Sep.:6393–6396.
28. Laumann A, et al. Development of a tongue-piercing method for use with assistive neuro-technology. *J. Amer. Med. Assoc.* submitted for publication.
29. Ergonomic requirements for office work with visual display terminals (VDTs)—Part 9: Requirements for non-keyboard input devices. *ISO 9241-9:2000(E).* 2002 Feb.
30. *Handbook of the International Phonetic Association.* Cambridge, MA, USA: Cambridge Univ. Press; 1999. International Phonetic Association
31. Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 1954; 47:381–91. [PubMed: 13174710]
32. Zhang X, MacKenzie IS. Evaluating eye tracking with ISO 9241—Part 9. *Proc. HCI Intl.* 2007 Jul.:779–788.
33. Williams MR, Kirsch RF. Evaluation of head orientation and neck muscle EMG signals as command inputs to a human—Computer interface for individuals with high tetraplegia. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2008 Oct; 16(10):485–496. [PubMed: 18990652]
34. Harada S, Landay JA, Malkin J, Li X, Bilmes JA. The vocal joystick: Evaluation of voice-based cursor control techniques. *Proc. ACM Conf. Computers and Accessibility—CHI.* 2006 Oct.:197–204.
35. McArthur V, Castellucci SJ, MacKenzie IS. An empirical comparison of Wiimote gun attachments for pointing tasks. *Proc. ACM Symp. Eng. Interactive Comp. Sys.* 2009 Jul.:203–208.
36. Soukoreff RW, MacKenzie IS. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *Int. J. Human-Computer Studies.* 2004 Dec.6:751–789.

Biographies

Xueliang Huo (S'07–M'12) 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, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, in 2011. His Ph.D. research was on developing wearable and wireless assistive technologies as well as human computer interfaces for people with severe disabilities.

He is currently a Hardware Engineer at Microsoft Corporation, Redmond, WA.

Hanguk Park (S'11) was born in 1980. He received both B.S. and M.S. degrees from Seoul National University, Seoul, Korea, in 2006 and 2008, respectively. He is currently working toward the Ph.D. degree in the GT-Bionics Laboratory, Georgia Institute of Technology, Atlanta,

From 2001 to 2004, he was with Bluebird-soft, where he designed circuit and system for industrial personal digital assistance (PDA). From 2008 to 2010, he worked for the Samsung-Electronics and designed High-Q RF bandpass filters for SAW-less receivers and PLL for cell-phone applications. He joined GT-Bionics Laboratory, Georgia Institute of Technology, in 2010. His research interests lies in system and IC design for biomedical applications, especially in Tongue Drive Systems.

Jeonghee Kim (S'11) was born in 1983. She received the B.S. degree in electrical engineering from the Kyungpook National University, Daegu, Korea, and the University of Texas at Dallas, Richardson, in 2007 and 2008, respectively, and the M.S. degree in electrical engineering and computer science at the University of Michigan, Ann Arbor, in 2009. She is currently working toward the Ph.D. degree at the Georgia Institute of Technology, Atlanta, in the GT-Bionics laboratory.

Her research interests are system design for biomedical devices with embedded mobile application, human computer interaction, and assistive technologies.

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.

From 2004 to 2007, he was an Assistant Professor in the Department of Electrical and Computer Engineering, NC State University, Raleigh, NC. In 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 110 peer-reviewed conference and journal publications.

Dr. Ghovanloo is an Associate Editor of the *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, *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 is the 2010 recipient of a CAREER award from the National Science Foundation. He is a member of the Tau Beta Pi, AAAS, Sigma Xi, and the IEEE Solid-State Circuits Society, IEEE Circuits and Systems Society, and IEEE Engineering in Medicine and Biology Society.

Fig. 1.
Top-level block diagram of the dTDS.

Fig. 2.
Major components of dTDS.

Fig. 3.
Block diagram of dTDS wireless headset.

Fig. 4.
dTDS experimental setup: a subject with SCI at level C4 wearing the dTDS headset and performing the typing task.

(a)

(b)

(c)

Fig. 5.
dTDS experimental GUIs: (a) maze navigation task, (b) typing task, (c) center-out tapping task [18].

Fig. 6.

Average completion time and typing rate of maze navigation experiment by able-bodied subjects using different devices.

Fig. 7.

Average total completion time of maze navigation experiment by SCI subjects using different devices.

Fig. 8.

Average completion time and typing rate in performing the center-out tapping task by able-bodied subjects using different devices.

Fig. 9.

Average throughput and error rate of different devices in completing center-out tapping tasks by able-bodied subjects.

Fig. 10.

Average total completion time of center-out tapping task by SCI subjects using different devices (error bars indicate the standard error considering the small number of subjects).

Fig. 11.

Seven naïve able-bodied subjects' response to a post-session questionnaire. Higher numbers represent more positive perceptions.

Fig. 12.
SCI patients' subjective rating on different devices at the end of sixth session.

TABLE I**Dual-Mode Tongue Drive System Hardware Specifications**

Specification	Value
<i>Magnetic Tracer</i>	
Material	NdFeB rare-earth magnet
Size (diameter and thickness)	Ø 3 mm × 1.6 mm
Residual magnetic strength	14500 Gauss
<i>Magnetic Sensors</i>	
Type	Aichi Steel AMI306 MI sensor
Dimensions	2.0 × 2.0 × 1.0 mm ³
Sensitivity / range	600 LSB/Gauss/±300 µT
<i>Microphone</i>	
Type	SiSonic SPM0408HE5H
Dimensions	4.7 × 3.8 × 1.1 mm ³
Sensitivity / SNR	−22 dB / 59 dB
<i>Control Unit</i>	
Microcontroller	TI-CC2510 SoC
Wireless frequency / data rate	2.4 GHz / 500 kbps
Sampling rate	50 sample/s/sensor
Number of sensors /duty cycle	4 / 8%
Audio codec / interface	TLV320AIC3204 / I2S
Audio sampling rate / resolution / compression	8 ksps / 16 bits / µ-Law
Operating voltage / total current	3 V / 35 mA (audio on) 3 V / 6 mA (audio off)
Dimensions	36 × 16 mm ²
<i>Headset</i>	
Rapid prototyping material	Objet VeroGray resin
Total weight	90 g (including battery)

TABLE II

Cursor Navigation And Typing Errors Using Different Devices In Completing Maze Navigation Tasks By Able-Bodied Subjects

	Devices			
	IMS/KB*	TDS	Dragon	dTDS
Navigation Error (Pixels)	0.12 ± 0.14	4.39 ± 2.94	3.99 ± 2.89	2.97 ± 2.56
Typing Error (%)	$0.7\% \pm 1.2\%$	$1.3\% \pm 2.1\%$	$6.2\% \pm 3.7\%$	$4.0\% \pm 3.5\%$

* MS/KB: Mouse/Keyboard

TABLE III

Effect of Different Factors on DTDS Maze Navigation Performance

Performance measures	Total Time (s)	Navigation Time (s)	Typing Rate (chars/min)	Navigation Error (pixel)	Typing Error (%)
<i>With TDS experience</i>	82.7± 9.0	33.4± 5.4	43.4± 8.7	2.0± 2.0	4.7± 4.2
<i>No TDS experience</i>	101.2 ± 23.2	44.6 ± 18.5	40.2 ± 11.5	4.0 ± 2.8	3.4 ± 3.0
<i>Native Speaker</i>	97.6 ± 14.4	42.6 ± 16.4	40.9 ± 10.0	4.3 ± 3.0	2.6 ± 2.9
<i>Non-native Speaker</i>	86.3 ± 23.1	35.4 ± 12.2	42.7 ± 10.5	1.6 ± 1.1	5.5 ± 3.8
<i>TDS Experience</i>	Sig. Power	0.017* 0.678	0.046* 0.521	0.248 0.208	0.984 0.050
<i>Accent</i>	Sig. Power	0.534 0.094	0.542 0.092	0.022* 0.645	0.222 0.218
<i>TDS Exp. * Accent</i>	Sig. Power	0.308 0.172	0.458 0.113	0.387 0.137	0.587 0.080

TABLE IV

Effect of Different Factors on DTDS Center-Out Tapping Performance

Performance measures	Total Time (s)	Navigation Time (s)	Typing Rate (chars/min)	Throughput (bits/s)	Error Rate (%)	Typing Error (%)
<i>With TDS experience</i>	95.8 ± 14.4	47.3 ± 8.6	43.0 ± 8.2	1.4 ± 0.3	19.9 ± 7.6	7.7 ± 5.7
<i>No TDS experience</i>	112.2 ± 19.5	59.9 ± 13.2	42.8 ± 12.1	1.1 ± 0.3	25.3 ± 12.6	5.1 ± 3.5
<i>Native Speaker</i>	104.2 ± 20.2	57.7 ± 13.9	47.4 ± 10.9	1.0 ± 0.3	23.8 ± 13.5	4.7 ± 4.2
<i>Non-native Speaker</i>	103.8 ± 18.4	49.5 ± 10.4	38.4 ± 7.1	1.4 ± 0.3	21.4 ± 6.9	8.1 ± 5.0
<i>TDS Sig</i>	0.003*	0.007*	0.120	0.04*	0.483	0.637
<i>Experience Power</i>	0.863	0.797	0.341	0.567	0.101	0.073
<i>Accent Sig.</i>	0.207	0.374	0.001*	0.966	0.792	0.356
<i>Power</i>	0.240	0.142	0.939	0.050	0.057	0.142
<i>TDS Exp. *</i>	0.735	0.497	0.395	0.191	0.628	0.632
<i>Accent Power</i>	0.063	0.103	0.134	0.245	0.074	0.073