Disability and Rehabilitation: Assistive Technology, 2011, 1–7, Early Online

ISSN 1748-3107 print/ISSN 1748-3115 online
DOI: 10.3109/17483107.2011.629330

informa
healthcare

RESEARCH ARTICLE

Evaluation of an ambient noise insensitive hum-based powered wheelchair controller

Tiago H. Falk[1,2], Alex Andrews[2], Fanny Hotzé[2], Eric Wan[2] & Tom Chau[2]

[1]*Institut National de la Recherche Scientifique, University of Quebec, Montreal, Quebec and and* [2]*Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, University of Toronto, Toronto, Ontario, Canada*

Purpose: **A recently-developed assistive technology nicknamed "the Hummer" was investigated as a potential powered wheelchair controller for individuals with severe and multiple disabilities. System performance in a noisy environment was compared to that obtained with a commercial automatic speech recognition (ASR) system.** ***Method:*** **A bi-hum driving protocol was developed to allow the Hummer to serve as a powered wheelchair controller. Participants performed several virtual wheelchair driving tasks of increasing difficulty using the two systems. Custom-written software recorded task execution time, number of commands issued and wall collisions, speed, and trajectory.** ***Results:*** **The bi-hum protocol was shown to be non-intuitive and required user training. Overall, the Hummer achieved lower performance relative to ASR. Once users became accustomed to the protocol, the difference in performance between the two systems became insignificant, particularly for the higher-difficulty task.** ***Conclusions:*** **The Hummer provides a promising new alternative for powered wheelchair control in everyday environments for individuals with severe and multiple disabilities who are able to hum, particularly for those with severe dysarthria which precludes ASR usage. A more intuitive driving protocol is still needed to reduce user frustration and mitigate user-generated errors; recommendations on how this can be achieved are given herein.**

Keywords: Ambient noise, assistive technology, automatic speech recognition, powered wheelchair control, vocal cord vibration

Implications for Rehabilitation

- The "Hummer" allows for noise-insensitive powered wheelchair control in everyday environments
- The "Hummer" accommodates individuals without functional speech or those with severe dysarthria to gain mobility independence, thus improving their quality of life.

Introduction

Today, approximately 6 million people in the United States are reported to be living with some kind of paralysis; 1.28 million of those are due to spinal cord injury (SCI) [1]. Of these individuals, 36% have reported "a lot of difficulty in mobility" and 16% as being "completely unable to move" [1]. Worldwide, as much as 33% of individuals with SCI are completely unable to move and need continuous care [2]. Technological, as well as medical progress, however, has drastically improved their quality of life, as well as allowed them to regain some independent mobility [1]. Representative technologies used for powered wheelchair control, for example, include tongue movement controllers (e.g. [3,4]), brain-computer interfaces [5], breath pressure [6] and sniffing [7] controllers, and speech recognition, either used alone (e.g. [8,9]) or in combination with humming for speed control [10].

Automatic speech recognition (ASR) has been researched for the last five decades, but only with the advances in computing power witnessed in the last decade, has ASR become mainstream and subsequently introduced in assistive devices [11]. Surveys with ASR users in the assistive technology realm (e.g. to use a computer), such as individuals with SCI, have suggested high satisfaction rates as well as positive perception of psychosocial impact [12,13]. When ASR is used for more safety-critical functions, however, such as controlling a powered wheelchair, satisfaction rates are much lower, with the majority of respondents using switch-operated systems as a backup [14].

Ambient noise (e.g. crowd noise, wind), user-generated noise (e.g. coughs, airways congestion, heavy breathing), and speech disorders (e.g. affecting as much as 50% of

Correspondence: Tiago H. Falk, Institut National de la Recherche Scientifique, University of Quebec 800 rue de la Gauchetiere West, Suite 6900, Montreal, Quebec, H5A-1K6, Canada. Tel: +1 (514) 228-7022. Fax: +1 (514) 875-0344. Email: tiago.falk@ieee.org

Canadian children with a disability [15]) are likely to be the culprits in such low satisfaction rates, as they are key ASR performance degrading factors. While there have been efforts to incorporate noise suppression and microphone array technologies into ASR-based wheelchair controllers, performance is still unacceptable for very noisy environments. As examples, average accuracies of 65% and 68% have been reported with single and multi-microphone systems after noise suppression in environments with 0 dB signal-to-noise ratio. For comparison, the same systems without noise suppression achieve average accuracies of 23% and 35%, respectively [16]. Throat microphones have also been used in the past as an additional layer of protection against ambient noise. Throat microphones, however, have been shown to cause errors due to user-generated noise sources, such as coughing, throat clearing, deep breathing, or even spastic head movements which cause the microphone to rub against the skin [17]. Furthermore, as evidence of the profound negative impact speech disorders have on ASR performance, recent studies have reported ASR accuracies to be around 4.5% for severely dysarthric speakers [18].

In order to overcome the limitations of voice-based technologies in everyday environments and for individuals with speech disorders, we have developed a non-intrusive vocal cord vibration monitoring device, nicknamed "the Hummer". The device focuses on key characteristics of the voice production system (i.e. periodic vocal cord vibrations) as opposed to the produced outcome (i.e. speech), which is vulnerable to ambient noise and vocal disorders, such as poor articulation [19]. The device was shown to be useful as a single-output assistive device [20,21] and, by discriminating between short- and long-duration vocalizations, as a dual-output device [22]. Motivated by the fact that phonatory control can be attained with individuals with severe voice disorders [23], this paper explores the use of the "Hummer" as an ambient noise insensitive hum-based wheelchair controller. By controlling the pitch frequency (i.e. frequency of vocal cord vibrations) of their hums, individuals are able to drive a powered wheelchair independently. More specifically, we have developed a bi-hum pipelined driving protocol where representative commands can include the use of rising and falling pitch frequencies to make the wheelchair go forward and backwards, respectively.

Materials and methods

Participants

Ten able-bodied participants (7 female, age: 25.4 ± 3.4 years) and one client participant with cerebral palsy and a mild-to-moderate speech disorder (female, age: 51) were recruited for the study. The study protocol was approved by the research ethics board of the hospital and all participants freely consented to participate. Participants had no history of using the Hummer or ASR for daily activities. The client participant had over 20 years of experience using a walker and a scooter but had no experience with powered wheelchair control.

Hummer: technology development

The technology behind the Hummer has been described in detail elsewhere [19], thus, only a brief description is provided here. It is known that speech sounds are produced by forced air from the lungs as it passes between the vocal cords. Voiced sounds (e.g. vowels) and hums, for example, cause periodic vibration of the vocal cords, with the frequency of vibration representing the person's pitch or tone. User-generated noises, on the other hand, have been shown to cause aperiodic vibrations [19]. By using a dual-axis accelerometer placed on the anterior surface of the throat (with axes of acceleration aligned to the anterior-posterior and superior-inferior anatomical axes) we can monitor vocal cord vibrations in real-time.

In previous studies aimed at the development of a single-output assistive device, we proposed the use of a normalized cross-correlation (NCC) function and NCC thresholding to detect periodicity in the vibratory signals [19–21]. Here, we are interested in not only detecting periodicity in the vibratory signals, but also the frequency at which they are vibrating. Due to properties of the NCC function, the distance between peaks is directly related to vibration frequency. For each user, a calibration session is needed in order to establish the user's minimum (f_{min}) and maximum (f_{max}) producible pitch frequencies. Different users have different ranges, with females generally producing higher frequencies. Once the pitch range has been established, thresholds can be set to discriminate between N distinct pitch levels. We have opted to perform bipartite discrimination (low and high) as research has shown that this is easier for individuals with dysarthria [23]. With the bipartite partition, pitch frequencies greater than the mid-range frequency $\left[\{f_{max}\text{-}f_{min}\}/2\right]$ are termed "high" and those lower are termed "low."

In the past, a prototype was developed using a PIC microcontroller; the device was portable and ran on either USB power or on two AA batteries [19]. Here, vocal cord vibrations are monitored in real-time using a USB bus-powered data acquisition device (National Instruments, Texas, USA) connected to a laptop PC running custom signal processing software written in Visual C++ (Microsoft, Washington, USA) with a National Instruments DAQmx application programming interface (National Instruments, Austin, TX, USA). Data were sampled at 1000 Hz with 32 bits per sample.

Driving protocol

Since it is difficult for individuals with pulmonary dysfunction or severe speech disorders to continuously increase/decrease their pitch levels, we developed a pipelined "bi-hum" protocol where users produce two consecutive hums of either varying or equal pitch frequency levels (i.e. high or low). In its current version, the protocol allows for up to one second between hums (this is configurable); beyond one second, the pipeline is cleared and the user must reinitiate the bi-hum process. If a low-frequency hum is followed by a high-frequency one, the "forward command" is activated. The opposite sequence (high-low) is used to activate the backward command. In order to control speed, the following commands were adopted. Two consecutive forward commands would increase wheelchair speed and three consecutive "forwards" would take the wheelchair to cruising speed. Once in cruise mode, a backward command can be issued to decrease speed and two "backwards" to stop the wheelchair. Alternately, a long-duration hum (1.5 seconds in the current version) of any frequency level can be used for an emergency stop. In consultation with potential users of the system, high-high and low-low pitch hum sequences were adopted to make the wheelchair go right and left, respectively (also configurable).

ASR baseline system

To gauge the usefulness of the hum-based wheelchair controller, we compared its performance with that obtained with a state-of-the-art commercial ASR system. In our experiments, Dragon Naturally Speaking 10 (Nuance, Massachusetts, USA) was used with an Audio DSP 400 USB close-talking microphone with noise cancelling capability (Plantronics, California, USA). Each user set up an individual profile and completed the calibration (training) session which required reading a short text. Settings were kept consistent for all participants (i.e. base vocabulary: general – large; language: US English; speech model: BestMatch III). The verbal commands used were: up (i.e. to go forward), down (i.e. to go backwards), left, right, and stop.

Experimental setup

A virtual wheelchair driving game was developed in Visual Basic.NET (Microsoft, Washington, USA) for the purpose of this study. The game consisted of three mazes of increasing difficulty levels; a snapshot of the three mazes in increasing difficulty can be seen in Figs. 1a–c, respectively. The aim of the game was to drive the wheelchair (blue circle) through the hallways (dark gray area) and come to a complete stop inside the room (cyan square). The yellow dot on the wheelchair represented the front of the wheelchair. The time to complete the task was recorded by the software, as were the number of collisions, number of commands issued, speed, and the trajectory taken. Each experimental session consisted of the participants concluding the three mazes using both the Hummer and the ASR baseline system. Each participant performed two sessions on two separate days; the system that was used first in a session was randomly chosen to reduce any biases in performance. Experiments were conducted in the presence of babble noise played via loudspeakers to investigate the usefulness of the proposed solution in everyday noisy environments. A Scosche SPL1000 digital sound level meter (Scosche, California, USA) was used to record noise levels, which ranged between 75–80 dB.

Performance metrics and statistical analysis

We use three performance metrics to gauge system performance: 1) task execution time, 2) number of commands (e.g. up, down, left, right, stop) issued and 3) number of wall collisions recorded during the completion of each maze. For comparison between the Hummer and the ASR baseline system, a repeated measures 3-way ANOVA considering session number, control device, and maze level was performed using SPSS.

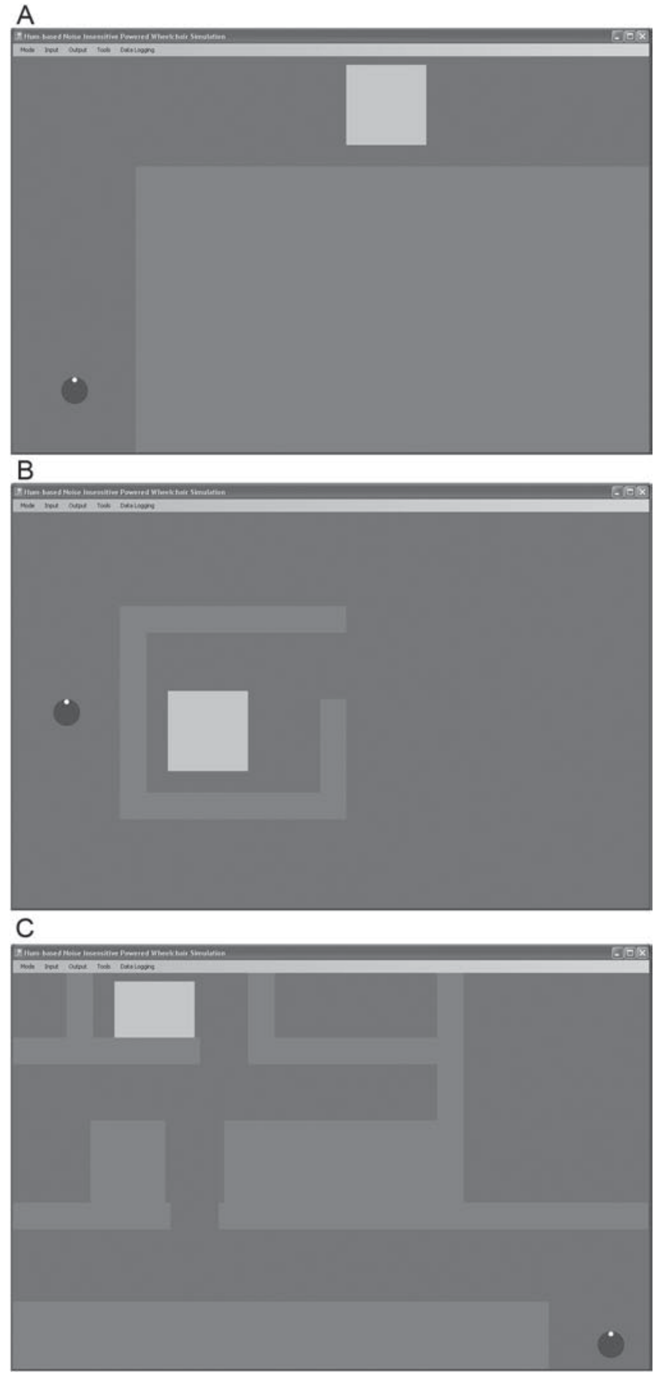

Figure 1. Snapshot of the three maze levels in the wheelchair driving program: (a) easy, (b) medium, and (c) difficult.

Results

Figs. 2a–c depict the average task execution time, number of commands issued, and number of wall collisions for both the Hummer and baseline ASR systems, respectively, during the first and second experiment sessions for the ten able-bodied participants. Client data were not included as she did not conclude the second experimental session due to personal reasons. For the Hummer, there was a significant improvement in all three performance metrics from the first to the second sessions. Notwithstanding, the ASR engine outperformed the Hummer in all metrics used. The ANOVA analysis reported

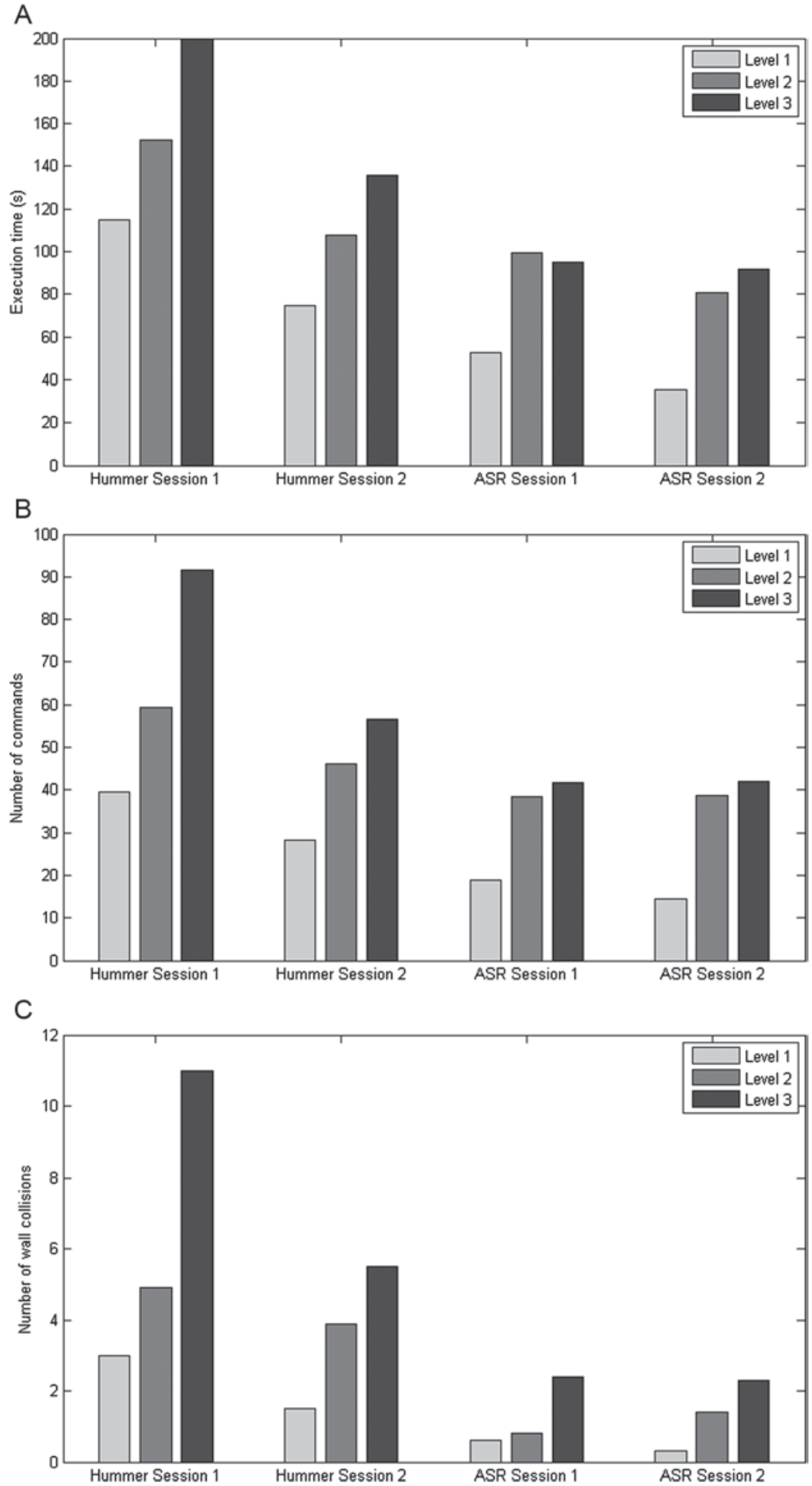

Figure 2. Performance comparison between the Hummer and ASR systems for the two experimental sessions for metrics (a) execution time, (b) number of commands issued, and (c) number of wall collisions across the three maze levels.

in Table 1 suggests that performance gains obtained with the ASR system were significant relative to those obtained with the Hummer (see main effects in Table I). Significant two-way interaction was also observed between session number and control device for all three performance metrics and between control device and maze level for the number of collisions metric. The three-way interaction effect was not significant.

Discussion

This study explored the use of a recently-developed assistive technology named "the Hummer" as a potential control device for powered wheelchairs. Performance comparisons with a commercial speech recognition system have suggested that updates to the device are still needed in order to make its use more intuitive for wheelchair control. More specifically, participants

Table I. Repeated measures 3-way ANOVA between Hummer and ASR control systems for the three performance metrics obtained from able-bodied participants

	p value		
Effect	Execution time	Number of commands	Number of collisions
Session	0.006	0.014	0.025
Control device	0.000	0.000	0.001
Maze level	0.000	0.000	0.002
Session × device	0.041	0.049	0.032
Session × level	0.593	0.33	0.174
Device × level	0.194	0.158	0.016
Session × device × level	0.16	0.126	0.282

emphasized the need for a more intuitive driving protocol as the pipelined bi-hum protocol was difficult to follow and caused a large number of erroneous activations. Users highlighted the fact that when a mistake was made in the first of the two hums (e.g. user generated a wrong pitch class), a method for resetting the erroneous command was not available. In lieu, users had to wait for the command to be reset automatically. This caused participants to lose their sense of control and led to mounting frustration and loss of confidence, which in turn, generated other mistakes, thus increasing the task execution time.

This precipitation of errors can be observed in the trajectory plot depicted by Fig. 3a. The top-left plot depicts the trajectory taken by Participant #7 during the maze level 3 (difficult) using the Hummer. Orange trajectories indicate slow speed; increasing speed is indicated by yellow and green (cruise speed) trajectories (the reader is referred to the online coloured version of this paper). A command issue is depicted by a "*" and a wall collision by a circle. As can be seen, an erroneous second turn (user erroneously activated a left turn as opposed to a right turn) caused a chain of wall collisions and command issues which, in turn, increased user frustration and contributed to the increase of the task execution time, number of commands issued, and yet further wall collisions. The top-right plot, in contrast, depicts the trajectory taken by the same participant but using the ASR system. Since the ASR requires only a single utterance command, such user-dependent mistakes can be avoided. Notwithstanding, environmental noise can cause ASR errors which may lead to wall collisions and an increased number of commands issued, as was the case depicted in the bottom-right plot of Fig. 3a for the second experiment session.

In order to overcome the limitation with the Hummer device, an alternate driving protocol may be implemented. Participants suggested the use of an ascending hum pitch for left turn, a descending hum pitch for right turn, a short duration hum (of any pitch frequency) for stop/reverse, and a long duration hum of constant pitch for forward movement. Generation of long-duration hums, however, may be an issue for individuals who are ventilated or those with poor pulmonary control [21,22]. One possible approach to overcome this limitation is to impose user-specific hum duration thresholds to discriminate between short- and long-duration hums. Our previous computer access experiments with the Hummer have suggested that the use of lower duration thresholds can be useful for individuals with poor pulmonary control [22].

An additional disadvantage of the bi-hum pipelined protocol arose when users had to generate a sequence of commands (e.g. two or three "forwards" in order to get to cruising speed). Feedback given by the participants suggested that it was not clear when a command pair had been executed and the system was ready to receive a second bi-hum command; such confusion caused erroneous commands to be generated. An example of such user-generated errors can be seen in the bottom-left plot of Fig. 3b. Near the end of the task, the participant attempted a sequence of commands to turn, slow down, and then stop the wheelchair in the finish room. Confusion about the temporal boundary between successive commands caused the user to erroneously activate the cruise speed command instead, causing her to cruise past the finish room. To correct this mistake, a large number of additional commands were needed as the user became frustrated and lost control of the wheelchair. In order to avoid such "feedback" errors, we are experimenting with a visual display marker which becomes blue when the system is ready to receive a new command (i.e. hum pair), amber when it has received the first hum and is awaiting the second, and green when it is executing a command. Such visual display will likely mitigate errors due to a lack of system feedback.

While overall the Hummer obtained significantly lower performance relative to the ASR system, significant improvement was obtained in Hummer performance from one experimental session to the next, suggesting that issues with the Hummer were related mostly to user training. The top- and bottom-left plots in Fig. 3a exemplify the performance improvement from the first to the second experimental session, respectively, for one of the participants. Similar trends were seen for the majority of the participants, thus corroborating findings shown in Fig. 2a–c. A statistical t-test showed that the gains obtained with training were significant for the Hummer for all three performance metrics (execution time: $p = 0.0134$, commands issued: $p = 0.0081$, and wall collisions: $p = 0.0351$). Since users were already accustomed to using speech on a daily basis, no such gains were seen with the ASR system (see Figure 2). In fact, the performance obtained with the Hummer during the second session was only significantly different from those obtained with the ASR system for maze level three (t-test, execution time: $p = 0.0154$; commands issued: $p = 0.0106$; wall collisions: $p = 0.0252$). Such findings suggest that training is an important factor if the Hummer is to be used as a wheelchair controller, particularly with the proposed driving protocol. Additional sessions with the Hummer are likely to further improve user performance.

Despite the lower performance obtained in this study, the Hummer presents two key advantages over ASR. First, the Hummer is insensitive to ambient noise and thus can be used in everyday environments. While ASR performance can be severely compromised by ambient noises, we have used a close-talking microphone with noise suppression capabilities to ameliorate these degrading factors. Noise suppression, however, is not very effective for wind noise as it is nonstationary

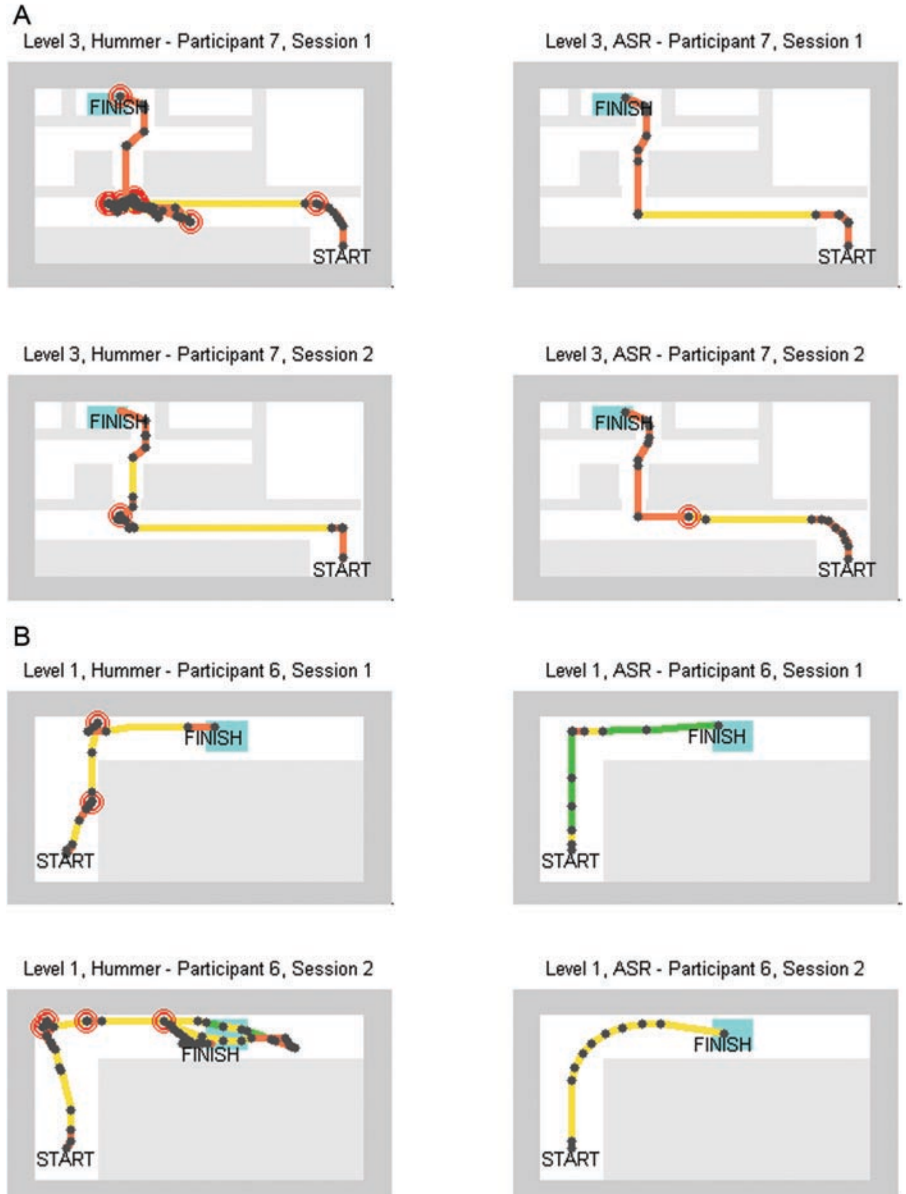

Figure 3. Trajectory plots for two able-bodied participants (subplots (a) and (b)) using the Hummer (left column) and the baseline ASR system (right column). The top row is for the first experimental session and the bottom for the second session. Wheelchair speeds are colour-coded, such that an orange trajectory indicates slow speed, yellow indicates medium speed and green cruising speed (readers are referred to the online coloured version of this paper). Command issues are depicted by a "*" and a wall collision by a circle.

and unpredictable [24]. As a consequence, the performance of ASR-controlled powered wheelchairs in open outdoor environments may be severely affected, even if throat microphones are used [25]. Future efforts should focus on quantifying the detrimental effects of wind noise on ASR performance for powered wheelchair control. Second, the Hummer accommodates individuals without functional speech or those with severe dysarthria to control a wheelchair independently. As mentioned previously, ASR accuracy can drop to 4.5% for severely dysarthric speakers [18], thus seriously compromising the usage of an ASR system for wheelchair control. As an example, Figure 4 depicts the trajectory plots obtained from the client participant. Evidently, ASR performance is severely compromised with dysarthric speech (successive command issues, illustrated by a "*" in the figure, show that the ASR system was unable to correctly recognize the user's commands). Unfortunately, due to personal reasons, the client participant was unable to conclude her second experimental session, thus we were not able to assess whether training had any positive effects on Hummer performance. In summary, the Hummer provides a promising alternative to wheelchair control in everyday environments, particularly for the individual with compromised speech. Updates to the driving protocol and user training, however, are needed before the Hummer can be effectively used as a powered wheelchair controller.

Acknowledgements

The authors wish to thank Ka Lun Tam, Steven Xu, and Pierre Duez for hardware and software support. This research was

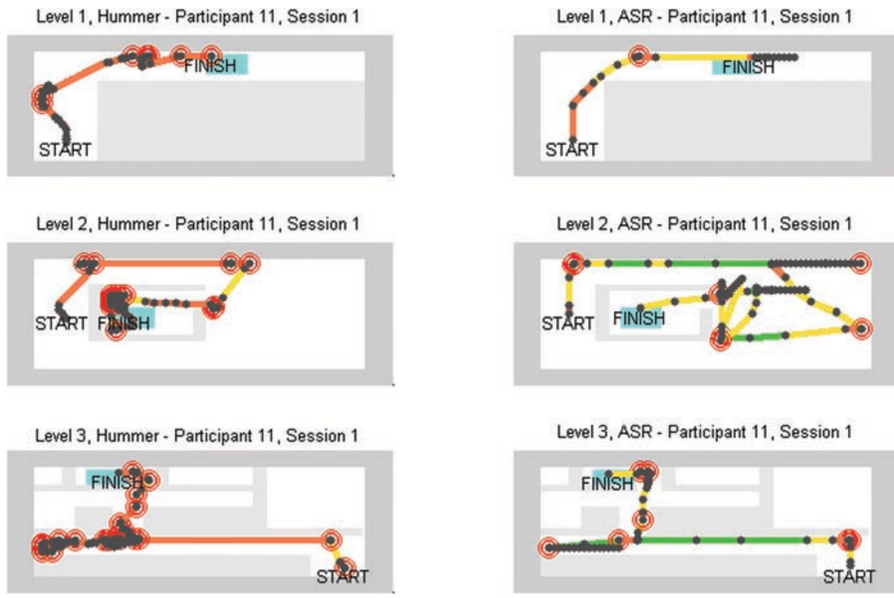

Figure 4. Trajectory plots for the client participant using the Hummer (left column) and the baseline ASR system (right column). The rows (from top to bottom) correspond to levels of increasing difficulty. Wheelchair speeds are colour-coded, such that an orange trajectory indicates slow speed, yellow indicates medium speed and green cruising speed (readers are referred to the online coloured version of this paper). Command issues are depicted by a "*" and a wall collision by a circle.

funded by the Natural Sciences and Engineering Research Council of Canada and their Canada Research Chair Program.

Declaration of Interest: The authors report no conflict of interest.

References

1. Christopher and Dana Reeve Foundation. One degree of separation: Paralysis and spinal cord injury in the United States, 2009. [cited, 2011, March 16]:Available at:http://www.christopherreeve.org/site/c.ddJFKRNoFiG/b.5091685/k.58BD/One_Degree_of_Separation.htm
2. Wyndaele M, Wyndaele JJ. Incidence, prevalence and epidemiology of spinal cord injury: what learns a worldwide literature survey? *Spinal Cord* 2006;44:523–529.
3. Huo X, Ghovanloo M. Evaluation of a wireless wearable tongue–computer interface by individuals with high-level spinal cord injuries. *J Neural Eng* 2010;7:12.
4. Lund ME, Christiensen HV, Caltenco HA, Lontis ER, Bentsen B, Andreasen Struijk LN. Inductive tongue control of powered wheelchairs. *Conf Proc IEEE Eng Med Biol Soc* 2010;2010:3361–3364.
5. Rebsamen B, Burdet E, Teo C, Zeng Q, Guan C, Ang M, Laugier C. A brain control wheelchair with a P300 based BCI and a path following controller, in Proceedings of the International Conference on Biomedical Robotics and Biomechatronics; 2006. pp 1101–1106.
6. Yamamoto M, Ikeda T, Sasaki Y. Real-time analog input device using breath pressure for the operation of powered wheelchair, in Proceedings of the International Conference on Robotics and Automation; 2008. pp 3914–3919.
7. Plotkin A, Sela L, Weissbrod A, Kahana R, Haviv L, Yeshurun Y, Soroker N, Sobel N. Sniffing enables communication and environmental control for the severely disabled. *Proc Natl Acad Sci USA* 2010;107:14413–14418.
8. Al-Rousan M, Assaleh K. A wavelet- and neural network-based voice system for a smart wheelchair control. *J Franklin Inst* 2011;348: 90–100.
9. Simpson RC, Levine SP. Voice control of a powered wheelchair. *IEEE Trans Neural Syst Rehabil Eng* 2002;10:122–125.
10. Nik H. Hum-Power controller for owered wheelchairs, MSc thesis, George Mason University, 2009.
11. Furui S. History and development of speech recognition. In Chen F, Jokinen K editors. *Speech technology: theory and applications*. New York, NY: Springer; 2010. pp 1–18.
12. Derosier R, Farber RS. Speech recognition software as an assistive device: a pilot study of user satisfaction and psychosocial impact. *Work* 2005;25:125–134.
13. Sears A, Karat C, Oseitutu K, Karimullah A, Feng J. Productivity, satisfaction, and interaction strategies of individuals with spinal cord injuries and traditional users interacting with speech recognition software. *UAIS* 2001;1:4–15.
14. Judge S, Robertson Z, Hawley M, Enderby P. Speech-driven environmental control systems–a qualitative analysis of users' perceptions. *Disabil Rehabil Assist Technol* 2009;4:151–157.
15. Statistics Canada, Participation and activity limitation survey 2006: Analytical report, Technical report, Ministry of Industry, Social and Aboriginal Statistics Division, Ottawa, ON (December, 2007).
16. Sasou A, Kojima H. Noise robust speech recognition applied to voice-driven wheelchair. *EURASIP J ADV SIG PR* 2009:8.
17. Lancioni G, Lems S. Using a microswitch for vocalization responses with persons with multiple disabilities. *Disabil Rehabil* 2001;17:209–221.
18. Rudzicz F. Towards a noisy-channel model of dysarthria in speech recognition, in Proceedings of the First Workshop on Speech and Language Processing for Assistive Technologies; 2010. pp. 80–88.
19. Falk TH, Chan J, Duez P, Teachman G, Chau T. Augmentative communication based on realtime vocal cord vibration detection. *IEEE Trans Neural Syst Rehabil Eng* 2010;18:159–163.
20. Chan J, Falk TH, Teachman G, Morin-McKee J, Chau T. Evaluation of a non-invasive vocal cord vibration switch as an alternative access pathway for an individual with hypotonic cerebral palsy – a case study. *Disabil Rehabil Assist Technol* 2010;5:69–78.
21. Lu E, Falk T, Teachman G, Chau T. Assessing the viability of a vocal cord vibration switch for four children with multiple disabilities *Open Rehabil J*, 2010; 3:55–61.
22. Lui M, Falk T, Chau T. Development and assessment of a dual-output vocal cord vibration switch for persons with multiple disabilities. *Disabil Rehabil Assist Technol* 2011, In Press.
23. Patel R. Phonatory control in adults with cerebral palsy and severe dysarthria. *Augment Altern Commun*. 2002;18:2–10.
24. Schmidt M, Larsen J, Hsiao F-T. Wind noise reduction using non-negative sparse coding, in Proceedings of the IEEE Workshop on Machine Learning for Signal Processing; 2007. pp. 431–436.
25. Winkler T, Pronkine S, Bardeli R, Köhler. J. A study of throat microphone performance in automatic speech recognition in motorcycles, in Proceedings of the NAG/DAGA Conference; 2009. pp. 1659–1662.