

Speech Biofeedback on Google Glass for People with Neuromotor Speech Impairments

Mansoor Pervaiz, Rupal Patel

Northeastern University, Boston, MA

mansoor@ccs.neu.edu, r.patel@neu.edu

Abstract

Individuals with neuromotor speech disorders due to conditions such as Multiple Sclerosis, Parkinson Disease and Cerebral Palsy have soft and slurred speech. These individuals require intervention to increase vocal loudness and to reduce speech rate. Traditional therapy is successful in a clinical setting but does not readily generalize to daily conversations. Technological innovations may help address the need for in-the-moment feedback. In previous work, we developed SpeechOmeter, a Google Glass application that allows users to monitor and increase their vocal loudness by providing visual feedback in real time. To incorporate speech rate feedback in the visual design, we modeled a set of user interface designs that simultaneously provide feedback about vocal loudness and speech rate. We then evaluated the efficacy and efficiency of these different designs and discovered a set of rules that may help to create unobtrusive visual feedback on small displays for individuals with cognitive and visual symptoms associated with neuromotor speech disorders.

Keywords

Neuromotor Speech disorders, Rehabilitation, User interfaces, Information interfaces and presentation, Usability evaluation, User centered design.

Introduction

Dysarthria is a speech impairment that accompanies neuromotor disorders such as Multiple Sclerosis (MS), Parkinson Disease (PD) and Cerebral Palsy (CP). It impacts strength, speed, range, steadiness, tone, and accuracy in one or more of the speech subsystems: respiration, phonation, resonance, and articulation (Duffy; Kathryn M. Yorkston et al.). Although different speech disturbances are found across dysarthria types, a common feature is reduced speech clarity (Duffy). Decreased intelligibility can lead to reduced communicative participation (Baylor et al.), defined as “taking part in life situations where knowledge, information, ideas, and feelings are exchanged” (Eadie et al. 309). Many patients consider the impaired ability to communicate to be one of the most difficult aspects of having dysarthria. Soft voice, monotone, breathy, hoarse voice quality, and imprecise articulation (Darley, Aronson, and Brown, “Differential Diagnostic Patterns of Dysarthria.”; Darley, Aronson, and Brown, “Clusters of Deviant Speech Dimensions in the Dysarthrias.”; Logemann et al.) contribute to limitations in communication in the vast majority of individuals (Pitcairn et al., “Non-Verbal Cues in the Self-Presentation of Parkinsonian Patients.”; Pitcairn et al., “Impressions of Parkinsonian Patients from Their Recorded Voices.”). Therefore, many speech interventions aim to increase speech intelligibility to enhance social participation and life fulfillment (Eadie et al.; O’Halloran, Hickson, and Worrall).

In a client-clinician setting, clinicians typically use two intervention strategies to increase speech intelligibility: 1) reducing speech rate (Kathryn M. Yorkston et al.; Yorkston and Beukelman; K. M. Yorkston et al.) and 2) increasing speech loudness (Solomon, McKee, and Garcia-Barry; Ramig et al.; Theodoros et al.). Lee Silverman Voice Treatment (LSVT), which focuses on speech loudness, has shown to improve speech clarity for individuals with PD (Ramig

et al.) and has also shown promise for individuals with MS (Sapir et al.), CP (Fox and Boliek), and Down Syndrome (Petska et al.). These treatments have limited efficacy because 1) clients have difficulty adhering to treatment in the absence of clinician cues, 2) LSVT requires maximal effort at all times which can be fatiguing, and 3) clinicians have limited insight into treatment adherence and performance.

SpeechOmeter

SpeechOmeter (Pervaiz and Patel) is a heads-up interface that provides users with real-time visual biofeedback cues regarding their vocal loudness when engaged in daily conversation. It allows users to speak at less than maximal effort by measuring the ambient noise level and motivating users to speak up to 5 dB higher by displaying the vocal loudness on the heads-up display. The display represents the target loudness range of 5 dB as 8 bars such that when a user speaks at 5 dB above the ambient level, all bars turn green. It also provides users and clinicians with performance and adherence statistics that can be computed over a few minutes to a full week so that clinicians can customize training.

Feedback Design

Inclusion of speech rate feedback required modification to the SpeechOMeter interface. The original SpeechOmeter interface only provided loudness cues in the form of a stack of 8 bars that turn green in proportion to the user's loudness level. The dual loudness and rate feedback display needed to account for cognitive and visual limitations associated with neuromotor disorders such as decreased information processing speed and diminished attention (Chiaravalloti and DeLuca; Berry). As a first step, we designed four feedback schemes and evaluated the efficacy of each on the target population.

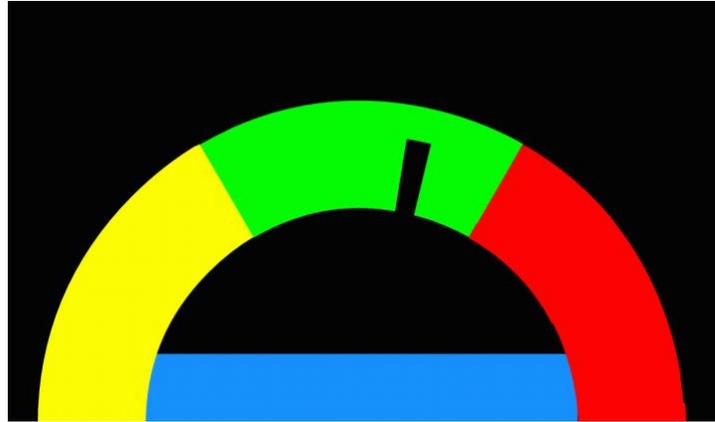


Fig 1. Feedback-1 illustrating speech rate with a thick solid arc and a needle sliding across from left to right and speech loudness by height of the colored fill underneath the arc.

The first scheme (Feedback-1) in Fig.1, illustrated speech rate in the form of a speedometer: a thick solid arc (180° degrees) with a needle slid across from left to right. The arc was divided into three equal regions of different color to show ideal and less than ideal speech rates. Speech loudness in Feedback-1 was represented by the height of a colored fill underneath the arc.

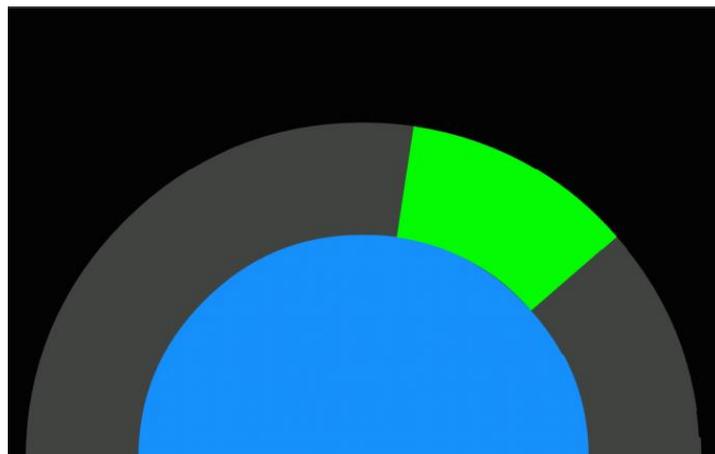


Fig 2. Feedback-2 illustrating speech rate with a 30 degree arc sliding over a curved path while speech loudness by height of the colored fill underneath the arc.

In Feedback-2 (Fig. 2), a 30° arc replaced the needle from the previous visual, slid over a curved path, and changed color when speech rate was too high. Meanwhile, the loudness visual remained unchanged from Feedback-1.

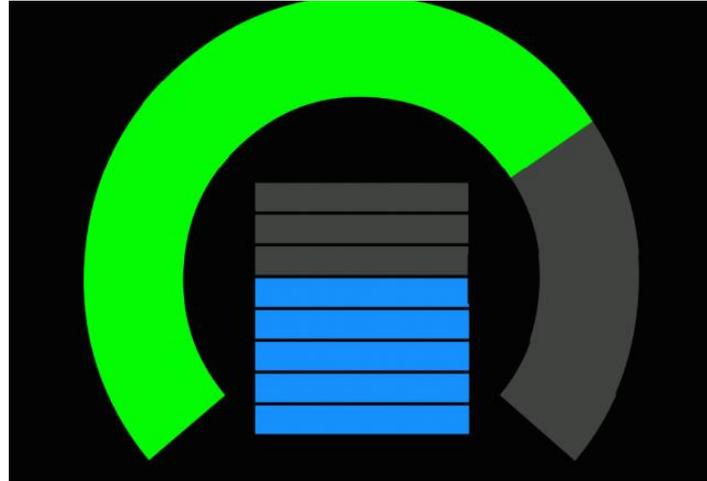


Fig 3. Feedback-3 illustrating speech rate with the fill of a 220 degree arc increasing from left to right and speech loudness by number of visible bars underneath the arc.

In Feedback-3 (Fig. 3), the fill of a 220° arc increased from left to right to show speech rate. Loudness was indicated by the number of visible bars, with up to 8 bars filled, that were stationed underneath the arc.



Fig 4. Feedback-4 illustrating speech rate with the number of visible steps of a discrete arc and speech loudness by the number of visible bars underneath the arc.

In Feedback-4 (Fig. 4) the arc from Feedback-3 was divided into 20 discrete steps and to show less than ideal speech rate it changed color beyond the 13th step (from the left). Loudness was again indicated by the number of visible bars, but the distance within the loudness bars was increased.

User Experiment 1

Five participants with speech disorders resulting from secondary progressive Multiple Sclerosis (MS) were recruited at The Boston Home, which is an assisted living facility for individuals with MS. Participation was voluntary with no compensation. Two participants were female and 3 male, with a mean age 58 years (range: 49 – 65). Participants were diagnosed between 20 and 40 years (mean: 31.6) previously.

Participants wore the Google Glass system in four 9-minute sessions. In each session, they viewed one feedback style and were asked to inform the experimenter if/when they noticed a change in speech rate or loudness feedback, where feedback was altered 10%, or 20%, or 30%. At the end of each session participants completed a SUS (System Usability Scale) questionnaire (Lewis and Sauro), a ten-item Likert scale that provided a global view of subjective assessments of usability.

Results

The results (Table 1) indicated that participants rated Feedback-1, with a needle sliding over an arc to represent speech rate, the highest (average score of 46.4 out of 50), and Feedback-4, with discrete speech rate and loudness, second highest (average score of 42.4) in terms of preference.

Table 1. Average SUS Scores

	Feedback 1	Feedback 2	Feedback 3	Feedback 4
Average SUS scores	46.4	31.6	39	42.4

During interviews, participants stated that Feedback-1 reminded them of a speedometer of an automobile. Individuals in the experiment favored Feedback-4 because it was easier to perceive changes when provided in discrete steps. Participants also appreciated the change in color when speech characteristics were not ideal.

Experiment Limitation

One limitation with this experiment was that participants continuously gazed at the feedback screen, which was not particularly naturalistic. Since typical conversations are more engaging, we modified the experiment to simulate a more realistic conversational interaction.

User Experiment 2

In Experiment 2, we documented the same 5 participants' input and gaze to the feedback screen during conversation for each of the four 9 minute sessions. To record participant input and gaze we developed two mobile applications: 1) User input recorder, and 2) User gaze monitor. The user input recorder allowed participants to tap on the phone screen every time they observed a change in the feedback. The user gaze monitor allowed researchers to monitor participant gaze by continuously touching the phone screen as long as the participant maintained their gaze to the feedback screen.

The Google Glass device displayed feedback to the participant who, instead of verbally informing the experimenter about the feedback change, recorded it through the user input

recorder. The experimenter documented the duration of a participant's gaze shift towards the Google Glass display on the user gaze monitor. During this experiment the researcher engaged the participant in conversation and asked questions to elicit continued dialogue. The number of misses (every time a participant missed a feedback change) and gaze duration towards the feedback screen were tabulated.

Results

The results (Table 2) showed that the percentage of misses and gaze percentage for each feedback were 86.2%, 73.7% (Feedback-1), 27.5%, 43.5% (Feedback-2), 24.5%, 41.7% (Feedback-3), and 17.7%, 38.2% (Feedback-4), which indicates that Feedback-4 had lower misses and lower gaze percentage.

Table 2. Average % misses for feedback change, and average % gaze towards UI screen

	Feedback 1	Feedback 2	Feedback 3	Feedback 4
% Misses	86.2%	27.5%	24.5%	17.7%
% Gaze	73.7%	43.5%	41.7%	38.2%

Discussion

In experiment-1, participants preferred Feedback-1 (Fig. 1) the most because it reminded them of an automobile speedometer, an entity that they are familiar with in real life. They also liked different color representations in the user interface for ideal, and less than ideal, performance. Participants found that it was easier to recognize changes in the feedback if the change was in discrete steps as opposed to continuous change.

Some of these results are not generalizable to daily conversation because during experiment-1 the participants continuously gazed at the feedback displayed by the Google Glass

device. This is not an acceptable behavior in a real conversation where conversational partners maintain eye contact. As a result, in the second experiment we elicited a conversation from the participants while changing the feedback. The participants were then asked to monitor visual feedback in their peripheral vision.

From experiment-2, we determined that despite the visual appeal of Feedback-1, this design performed poorly when participants were engaging in regular conversation. Participants missed a higher percentage of feedback changes, despite frequently gazing the UI screen. In contrast, participants were able to monitor feedback change better in Feedback-4 (Fig. 4) while spending less time gazing at the UI screen. Besides Feedback-4, Feedback-2 and Feedback-3 also performed better than Feedback-1. The reason Feedback-1 had higher error rate was because it is difficult to monitor a small needle in peripheral vision.

Conclusions

We applied user center design process to extend the SpeechOmeter interface to include speech rate feedback for individuals with neuromotor disorders. The extended system will provide training for two commonly used speech intervention strategies 1) increasing speech loudness and 2) reducing speech rate. In a daily conversation, users will be able to monitor speech loudness and to increase speech clarity. The ultimate goal is that these online visual cues will improve their social participation and life fulfillment.

With regards to interface design, the current user experiments suggest that visual feedback for users with neuromotor disorders and concomitant motor and visual impairments should be simple and easy to view. Performance changes in feedback should be in discrete steps and unsuccessful performances should be highlighted by a change in color. Additionally, for ecological validity, we learned that user experiments should be performed while engaged in

conversational dialogue.

Acknowledgements

We would like to thank the residents of The Boston Home and their speech therapist (Alex Burnham) for their feedback during the design process and participation in the studies.

Works Cited

- Baylor, Carolyn et al. "Variables Associated with Communicative Participation in People with Multiple Sclerosis: A Regression Analysis." *American Journal of Speech-Language Pathology* 19.2 (2010): 143–153. Print.
- Berry, E. "Slowing of Reaction Time in Parkinsons Disease: The involvement of the Frontal Lobes." *Neuropsychologia* 37.7 (1999): 787–795. Web.
- Chiaravalloti, Nancy D., and John DeLuca. "Cognitive Impairment in Multiple Sclerosis." *The Lancet Neurology* 7.12 (2008): 1139–1151. Web.
- Darley, F. L., A. E. Aronson, and J. R. Brown. "Clusters of Deviant Speech Dimensions in the Dysarthrias." *Journal of speech and hearing research* 12.3 (1969): 462–496. Print.
- . "Differential Diagnostic Patterns of Dysarthria." *J Speech Hear Res* 12.2 (1969): 246–269. Print.
- Duffy, Joseph R. *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*. Elsevier Health Sciences, 2012. Print.
- Eadie, Tanya L. et al. "Measuring Communicative Participation: A Review of Self-Report Instruments in Speech-Language Pathology." *American Journal of Speech-Language Pathology* 15.4 (2006): 307–320. Print.
- Fox, Cynthia M., and Carol A. Boliek. "Intensive Voice Treatment (LSVT LOUD) for Children with Spastic Cerebral Palsy and Dysarthria." *Journal of Speech, Language, and Hearing Research* 55.3 (2012): 930–945. Print.
- Lewis, James, and Jeff Sauro. *Human Centered Design*. Springer Berlin Heidelberg (2009): 94–103. Print.

- Logemann, J. A. et al. "Frequency and Cooccurrence of Vocal Tract Dysfunctions in the Speech of a Large Sample of Parkinson Patients." *The Journal of speech and hearing disorders* 43.1 (1978): 47–57. Print.
- O'Halloran, Robyn, Louise Hickson, and Linda Worrall. "Environmental Factors That Influence Communication between People with Communication Disability and Their Healthcare Providers in Hospital: A Review of the Literature within the International Classification of Functioning, Disability and Health (ICF) Framework." *International Journal of Language & Communication Disorders* 43.6 (2008): 601–632. Web.
- Pervaiz, Mansoor, and Rupal Patel. "SpeechOmeter: Heads-up Speech Monitoring to Improve Speech Clarity." *Proceedings of the 16th international ACM SIGACCESS conference on Computers and accessibility* (2014): n. pag. Web.
- Petska, J. et al. "LSVT® and Children with Down Syndrome: A Pilot Study." *Poster Session Presented at the 13 The Biennial Conference on Motor Speech, Austin, TX.* N.p., 2006. Print.
- Pitcairn, T. K. et al. "Impressions of Parkinsonian Patients from Their Recorded Voices." *The British journal of disorders of communication* 25.1 (1990): 85–92. Print.
- . "Non-Verbal Cues in the Self-Presentation of Parkinsonian Patients." *The British journal of clinical psychology / the British Psychological Society* 29 (Pt 2) (1990): 177–184. Print.
- Ramig, L. O. et al. "Intensive Speech Treatment for Patients with Parkinson's Disease: Short-and Long-Term Comparison of Two Techniques." *Neurology* 47.6 (1996): 1496–1504. Print.

Sapir, S. et al. “Effects of Intensive Phonatory-Respiratory Treatment (LSVT) on Voice in Two Individuals with Multiple Sclerosis.” *Journal of Medical Speech-Language Pathology* 9.2 (2001): 141–151. Print.

Solomon, Nancy P., Amy S. McKee, and Sandra Garcia-Barry. “Intensive Voice Treatment and Respiration Treatment for Hypokinetic-Spastic Dysarthria after Traumatic Brain Injury.” *American Journal of Speech-Language Pathology* 10.1 (2001): 51–64. Print.

Theodoros, D. G. et al. “The Effects of the Lee Silverman Voice Treatment Program on Motor Speech Function in Parkinson Disease Following Thalamotomy and Pallidotomy Surgery: A Case Study.” *Journal of Medical Speech-Language Pathology* 7.2 (1999): 157–160. Print.

Yorkston, Kathryn M. et al. “Evidence for Effectiveness of Treatment of Loudness, Rate, or Prosody in Dysarthria: A Systematic Review.” *Journal of Medical Speech-Language Pathology* 15.2 (2007): XI–XXXVI. Print.

Yorkston, K. M. et al. “The Effect of Rate Control on the Intelligibility and Naturalness of Dysarthric Speech.” *The Journal of speech and hearing disorders* 55.3 (1990): 550–560. Print.

Yorkston, K. M., and D. R. Beukelman. “Ataxic Dysarthria: Treatment Sequences Based on Intelligibility and Prosodic Considerations.” *The Journal of speech and hearing disorders* 46.4 (1981): 398–404. Print.