Treating Speech-Movement Hypokinesia in Parkinson's Disease: Does Movement Size Matter?

Elaine Kearney[1,2], Brandon Haworth[2,3], Jordan Scholl[1,2], Petros Faloutsos[2,3], Melanie Baljko[2,3], Yana Yunusova[1,2,4]

[1]Department of Speech-Language Pathology, University of Toronto, Toronto, Canada; [2]University Health Network – Toronto Rehabilitation Institute; [3]Department of Electrical Engineering & Computer Science, York University, Toronto, Canada; [4]Sunnybrook Research Institute, Biological Sciences, Toronto, Canada

Contact Author:

Elaine Kearney

elaine.kearney@mail.utoronto.ca

Department of Speech-Language Pathology

Rehabilitation Sciences Building

Faculty of Medicine, University of Toronto

160-500 University Avenue

Toronto, ON, Canada

M5G 1V7

Abstract

Purpose: This study evaluates the effects of a novel speech therapy program that uses a verbal cue and gamified augmented visual feedback regarding tongue movements to address articulatory hypokinesia during speech in individuals with Parkinson's disease (PD).

Method: Five participants with PD participated in an ABA single-subject design study. Treatment aimed to increase tongue movement size using a combination of a verbal cue and augmented visual feedback and was conducted in 10 45-minute sessions over five weeks. The presence of visual feedback was manipulated during treatment. Articulatory working space (AWS) of the tongue was the primary outcome measure and was examined during treatment as well as in cued and uncued sentences pre and post treatment. Changes in speech intelligibility in response to a verbal cue pre and post treatment were also examined.

Results: During treatment, 4/5 participants showed a beneficial effect of visual feedback on tongue AWS; at the end of the treatment they used larger tongue movements when cued, relative to their pre-treatment performance. None of the participants, however, generalized the effect to the uncued sentences. Speech intelligibility of cued sentences was judged as superior post-treatment only in a single participant.

Conclusions: This study demonstrated that using an augmented visual feedback approach is beneficial, beyond a verbal cue alone, in addressing articulatory hypokinesia in individuals with PD. An optimal degree of articulatory expansion might, however, be required to elicit a speech intelligibility benefit.

Keywords: Parkinson's disease, dysarthria, augmented visual feedback, tongue movements, speech games

Parkinson's disease (PD) is the second most common degenerative disease of neurological origin (Bertram & Tanzi, 2005), and nearly 90% of individuals with PD develop a motor speech disorder affecting phonatory, prosodic, and articulatory aspects of speech (Ho, Iansek, Marigliani, Bradshaw, & Gates, 1998). The existing treatment approaches for dysarthria in PD primarily focus on the phonatory/prosodic aspects of speech (Atkinson-Clement, Sadat, & Pinto, 2015). The most commonly prescribed treatment, the Lee Silverman Voice Treatment program (LSVT LOUD; Ramig, Countryman, Thompson, & Horii, 1995; Ramig et al., 2001), aims to address the reduced vocal loudness associated with hypokinetic dysarthria. Other approaches include the Pitch Limiting Voice Treatment (PLVT; de Swart, Willemse, Maassen, & Horstink, 2003), treatment using the SpeechVive device (Richardson, Sussman, Stathopoulos, & Huber, 2014), Speech Rate and Intonation Therapy (SPRINT; Martens et al., 2015), prosodic exercises (Scott & Caird, 1984), and rate reduction techniques (Lowit, Dobinson, Timmins, Howell, & Kröger, 2010). In addition to these treatments, less formalized approaches that use global therapy techniques – speaking louder or with increased clarity – are used with this population (Yorkston, Hakel, Beukelman, & Fager, 2007). Among these treatment methods, those directly targeting articulatory deficits, experienced by 45% of individuals with PD (Logemann, Fisher, Boshes, & Blonsky, 1978), remain limited.

The Articulatory Disorder in PD

Studies of articulatory movements in individuals with PD have indicated that hypokinesia and bradykinesia are observed in movements of the jaw (Darling & Huber, 2011; Forrest, Weismer, & Turner, 1989; Kearney et al., 2017; Walsh & Smith, 2012), lips (Ackermann, Konczak, & Hertrich, 1997) and, possibly, the much less studied tongue (Weismer, Yunusova, & Bunton, 2012; Yunusova, Weismer, Westbury, & Lindstrom, 2008). These findings have been reported in a variety of speech tasks from syllable repetitions to sentence and passage readings. There is an emerging literature linking changes in articulatory movements to speech intelligibility in PD. When reading a passage, slower movements of the tongue body, but not of the jaw or tongue tip, were correlated with reduced speech intelligibility (Weismer et al., 2012). However, only speakers with a mild dysarthria impairment were sampled in this study. Kearney et al. (2017) showed a positive association between speech intelligibility and articulatory movement size of the jaw, tongue blade, and tongue dorsum in a group of speakers with PD, who exhibited a range of speech intelligibility deficit. The smaller tongue blade movements were consistently – across various sentences – associated with more impaired intelligibility as compared to the jaw and tongue dorsum movements. We reasoned that a motor intervention focusing on an increase of the tongue blade movement size during speech would result in a direct effect on speech intelligibility.

Scaling Movement as a Treatment Approach for Hypokinesia in PD

The basal ganglia play an important role in the scaling and maintaining of movement amplitude in voluntary movement tasks (Desmurget, Grafton, Vindras, Grea, & Turner, 2004), which can help explain the underscaling of movement amplitude observed across both gross and fine motor tasks in PD, including balance, gait, handwriting, and speech. Importantly, patients with PD can modulate movement amplitudes, particularly when externally cued by an instruction or a visual prompt (Ford, Malone, Nyikos, Yelisetty, & Bickel, 2010; Oliveira, Gurd, Nixon, Marshall, & Passingham, 1997). Practicing larger amplitudes may, in turn, enhance activation of damaged neural pathways or slow their decline (Beall et al., 2013; Farley, Fox, Ramig, & McFarland, 2008). Cueing to increase movement amplitude is commonly used in motor therapy for improving gait length (Spaulding et al., 2013), reaching amplitude (Ebersbach et al., 2010) and size of handwriting (Nackaerts, Nieuwboer, & Farella, 2017). Such an approach may also be beneficial in speech therapy.

Existing speech interventions based on increasing vocal loudness, such as LSVT LOUD, are said to result in changes across multiple subsystems of speech including articulation, resulting in the upscaling of articulatory movements (Yorkston, Hakel, et al., 2007). It has been suggested that the larger articulatory movements may facilitate an improved distinction between speech sounds (Fox, Ebersbach, Ramig, & Sapir, 2012). Studies to-date examined the effect of a one-time loud-speech instruction on speech kinematics in individuals with PD and showed larger and faster movements of the jaw and tongue (Darling & Huber, 2011; Dromey, 2000; Kearney et al., 2017). To the best of our knowledge, these effects have not been examined pre and post treatment. Further, those with articulatory deficits (e.g., "significant rate disorders") show relatively poor outcomes post-LSVT LOUD (Fox et al., 2012, p. 7). These findings suggest that a direct articulatory intervention might be needed for speakers with articulatory hypokinesia as it might be beneficial in addressing this underlying pathophysiology. Further, a more targeted treatment approach may be less effortful for individuals with PD, who experience significant disease-related fatigue (Karlsen, Larsen, Tandberg, & Mæland, 1999).

Training speakers with hypokinesia to upscale their movement size may be conducted by cueing a patient with a simple verbal prompt regarding the relevant movement parameters. A verbal cue alone, however, would require individuals to rely on their own proprioception and sensorimotor integration to detect and implement changes in their articulatory movements. Deficits in both proprioception and sensorimotor integration are common in PD (Mollaei, Shiller, & Gracco, 2013; Schneider, Diamond, & Markham, 1986), and these deficits may make a verbal cue alone insufficient in modifying movement parameters. Further, it is challenging for a clinician to reliably judge spatial properties of orofacial movements in order to assess performance and provide feedback to an individual during therapy (Simione, Wilson, Yunusova, & Green, 2016). For these reasons, we chose to use the combination of a verbal cue and augmented visual feedback (AVF) to target articulatory hypokinesia in therapy.

Our group recently developed and tested the feasibility of an AVF system that provides information about articulatory movements and aims to remediate articulatory hypokinesia in individuals with PD (Haworth, Kearney, Baljko, Faloutsos, & Yunusova, 2014; Shtern, Haworth, Yunusova, Baljko, & Faloutsos, 2012; Yunusova et al., 2017). The system employs electromagnetic articulography via the Wave Speech Research System (Northern Digital Inc., Canada) and provides visual information regarding movements of a single sensor attached to the tongue blade. Tongue movement size is indexed by articulatory worksing space (AWS) – a global measure of articulatory movement size taken across an entire speech utterance (Kearney et al., 2017; Weismer et al., 2012). Yunusova et al.'s (2017) feasibility study showed that, following a single training session where tongue AWS during sentence productions was visualized in the form of a game, individuals with PD were able to increase articulatory movement size. Further, the effects of training were evident at a retention session 24 hours later. This system has not yet been examined in the context of a structured treatment program, and the improvements in hypokinesia have not been assessed with respect to changes in speech intelligibility.

AVF in Rehabilitation of Movement Disorders and Speech Production

AVF is an external source of feedback that can supplement an individual's own somatosensory and auditory feedback during motor skill learning (Swinnen, 1996). A recent systematic review by our group of AVF-aided motor interventions in PD – including 10 randomized control trials (RCTs) – showed that outcomes following treatment with AVF were often superior to those following traditional rehabilitation (Kearney, Shellikeri, Martino, & Yunusova, 2018). Eight of nine RCTs that provided raw data from which we could calculate effects sizes showed small to large effects in activity-level measures of balance and gait after training with AVF as compared to traditional therapy. Although these findings may not be directly applicable to speech motor control, they can provide a useful startpoint in the absence of speech-specific empirical data to guide treatment design (Grimme, Fuchs, Perrier, & Schöner, 2011).

To date, AVF has not been systematically studied in the rehabilitation of the speech movement disorder in PD. To the best of our knowledge, only one previous study examined the effect of AVF for improving vocal loudness in PD (Scott & Caird, 1984). The results were comparable between the experimental group and a control group who received similar treatment without visual feedback. The groups, however, were not randomly assigned, and the results are difficult to interpret due to baseline differences in the outcome measures between the experimental and control groups. Notably, AVF has been successfully applied for the remediation of speech movement disorders in other clinical populations, such as acquired apraxia of speech (Katz, Bharadwaj, & Carstens, 1999; McNeil et al., 2010) and speech sound disorder (Cleland, Scobbie, & Wrench, 2015; Dent, Gibbon, & Hardcastle, 1995). AVF interventions are of high interest currently due to the recent explosion in technology development (Campbell & Yunusova, 2017).

Current Research

We conducted a Phase 1 clinical-outcome research study to identify the therapeutic effects of a 10-session articulatory-treatment program using a verbal cue and AVF for individuals with PD (Robey, 2004). Given the articulatory nature of the intervention, tongue movement size - indexed by AWS - was the primary outcome measure and was evaluated in a series of analyses. First, we examined articulatory movements in three baseline sessions to assess the stability of AWS before treatment. Second, we examined the effect of a simple verbal cue on AWS before treatment to answer the question whether and to what extent a verbal cue alone resulted in changes in movement size. Third, we evaluated the effect of AVF (+ verbal cue) on AWS during treatment as compared to trials with the verbal cue alone in order to establish the direct effect of AVF on articulatory kinematics. Fourth, the effects of treatment were evaluated via the following analyses: (1) The effect of the cue on trained sentences pre and post treatment was examined to assess whether the treatment was effective in teaching participants to use the large movement cue; (2) The effect of the cue on the untrained sentences was assessed to determine the generalization of cueing from trained to untrained sentences; and (3) The generalization of treatment to untrained (uncued) sentences was assessed to judge whether the "large movement" strategy was habituated to novel sentences. Finally, changes in speech intelligibility for sentences produced with a verbal cue were examined pre-post treatment to address the question of whether increases in tongue movement size corresponded to improvements in speech intelligibility.

We hypothesized that the response to the verbal cue before treatment would be limited and that AVF (+ verbal cue) would result in a greater increase in tongue movement size during treatment compared to trials with verbal cues alone. Further, we expected that individuals with PD would be able to increase their tongue movement size in response to a verbal cue post treatment and that the effect would generalize to untrained cued sentences as well as uncued sentences. Finally, we hypothesized that increases in tongue movement size in response to a verbal cue would correspond to improvements in speech intelligibility.

Methods

Ethics

Ethical approval for this project was granted by the University Health Network–Toronto Rehabilitation Institute Research Ethics Board and the Health Sciences Research Ethics Board at the University of Toronto. All participants provided informed written consent before starting the study.

Participants

Five adults, diagnosed with PD and native speakers of English, were recruited from a larger parent study of speech kinematics and speech intelligibility in individuals with PD (N = 21, Kearney et al., 2017). Only participants who exhibited clear evidence of hypokinetic dysarthria with reduced articulatory movement size, based on the findings of the original kinematic study, and a speech intelligibility deficit were recruited for the present study. Exclusion criteria for the parent study were a history of other neurological disorders or conditions affecting speech as well as uncorrected vision impairment and hearing loss. A further exclusion criterion was the enrollment in speech therapy at the time of the study.

Demographic, clinical, and speech characteristics of the participants are provided in Table 1. All participants were male, with a mean age of 75.45 years (SD = 8.71). On average, participants were diagnosed with PD 3.14 years (SD = 1.55) before the study, and their Hoehn and Yahr scores were between 1 and 2, indicating a relatively early stage of the disease. Performance on the Montreal Cognitive Assessment (MoCA) indicated an absence of dementia (Nasreddine et al., 2005). All participants reported that they were on medication to alleviate PD-related symptoms, and their medications did not change for the duration of the study. The study sessions were scheduled when the participants felt optimally medicated and at a consistent time in their medication cycle throughout the study.

Previous history of speech therapy was negative in the majority of cases. One participant (PD30) reported previously attending some speech therapy sessions approximately four years before the study but was unable to recall the details of the therapy. All participants reported normal or corrected-to-normal vision. 4/5 had pure tone thresholds of 40dB or better in at least one ear at 1000, 2000, and 4000Hz (Ventry & Weinstein, 1983). One participant (PD25) presented with high-frequency hearing loss bilaterally[1] and wore hearing aids during the screening session; however, he did not wear the hearing aids for the remainder of the study due to potential interference with the electromagnetic recording equipment.

All participants in the current study showed intelligibility impairment at baseline on at least one of the intelligibility measures (SIT < 96% or scaled intelligibility *Z* score < -1.0, expressed relative to healthy control speakers). These baseline intelligibility ratings were obtained during the parent study (Kearney et al., 2017) using (1) a sentence transcription task (Sentence Intelligibility Test; Yorkston, Beukelman, Hakel, & Dorsey, 2007) and (2) a scaled intelligibility task (direct magnitude estimation (DME) with modulus; Weismer & Laures, 2002; Yunusova, Weismer, Kent, & Rusche, 2005). Perceptual characteristics of dysarthria were independently judged by two speech-language pathologists (SLPs). The SLPs listened to recordings of *My Grandfather* passage by each speaker and identified prominent deviant perceptual characteristics. The SLPs were provided with a list of the most commonly associated

[1] PD25 was recorded as part of the original kinematic study but his data were excluded from the group analysis. He was included in the current study as the analysis was conducted at the individual level.

perceptual characteristics of hypokinetic dysarthria (Darley, Aronson, & Brown, 1969) and were encouraged to record other perceptual characteristics if noted (McRae, Tjaden, & Schoonings, 2002). The perceptual ratings revealed that all participants presented with articulatory symptoms (e.g., imprecise consonants, change in rate) in addition to phonatory or prosodic symptoms (e.g., monoloudness, reduced stress). Perceptual characteristics identified by both SLPs are indicated in bold in Table 1.

Instrumentation and Signal Processing

Tongue movements were recorded during assessment and treatment sessions using the Wave (Northern Digital Inc., Canada), a 3D electromagnetic tracking system with sub-millimeter accuracy (Berry, 2011). Movement data were used to assess performance as well as to provide visual feedback during treatment. A 6 degree-of-freedom (DOF) sensor attached to a headband was placed on the forehead, and a 5-DOF sensor was attached to the tongue blade using non-toxic dental glue (PeriAcryl®90, Glustitch). The tongue sensor was placed at midline, approximately 10 millimeters from the tongue tip (mean distance across all participants/sessions = 10.26mm, *SD* = 1.54). Movement data were acquired at a sampling rate of 100Hz and were post-processed to subtract head movements and to filter the data using a median filter (window size = 3) in real time (Haworth, Kearney, Faloutsos, Baljko, & Yunusova, 2018). A 100Hz sampling rate adequately captured tongue movements – typically occurring below 15Hz during speech (Gracco, 1992) – and allowed for consistent data buffering from the recording computer to the visual feedback computer (Haworth et al., 2018).

To provide visual feedback, tongue movement data during each sentence recording were accessed using the Wave Real-Time Application Program Interface (RTAPI) and transferred to a visualization computer using the Wave proxy server. Visual feedback regarding tongue movement size (see Measurements section below) was provided in a game format using Unity3D v4.6.5p1 game engine technologies (Unity Technologies Inc, 2015). Games were displayed on a 24” 24-bit colour LCD monitor. Participants sat approximately 140cm from the monitor, positioned at eye-level.

Simultaneously, acoustic data were recorded via a lapel microphone (Countryman B3P4FF05B) positioned 15 cm from the participant’s mouth and digitized at 22kHz, 16-bit resolution on the hard drive of a computer.

Experimental Design

An ABA single-subject design was used to evaluate the effects of treatment for each participant in the study. Multiple baseline measures were taken during the first ‘A’ phase to establish the stability of articulatory performance before beginning treatment (Kazdin, 2011) and to evaluate the effect of a verbal cue alone. Additionally, data from the ‘B’ (treatment) phase was used to examine the effect of AVF on articulatory movements. Finally, measures taken pre and post treatment were used to examine the treatment and generalization effects on articulatory movements and to assess if changes in articulatory movements corresponded to changes in speech intelligibility.

Outcome Measures

Two measures were employed in this study. (1) Tongue AWS was chosen because in previous research it showed sensitivity to disease-related change in individuals with PD (Kearney et al., 2017; Weismer et al., 2012) and the effect of training (Yunusova et al., 2017). AWS was calculated as the volume of a convex hull fit to the movement trajectory of each sentence using a MatLab function *convhull*. Figure 4 shows an example AWS for a single sentence (“Jimmy worked on a crossword puzzle”) produced by PD28 pre and post treatment.

The measure is shown in two-dimensional (2D) space for simplification; however, the measurements were conducted in 3D.

(2) Paired-comparison ratings of speech intelligibility were used to assess if participants were easier to understand before or after treatment, or if there was no difference between the sentences produced at the two time points. The percentage of ratings categorized as being easier to understand pre or post treatment, or as being the same pre-post treatment, was calculated out of the total number of ratings per speaker (n = 20).

Assessment and Treatment Schedule and Procedures

All assessment and treatment sessions were conducted by the third author, who was not involved in the design of the study or analysis of the data. Figure 1 shows a flowchart of the assessment and treatment schedule. Following recruitment, participants attended three baseline assessment sessions, 10 treatment sessions, and a single post-treatment assessment session.

Pre-Treatment Assessment. During the first three baseline sessions, participants read six sentences, each repeated four times in a blocked fashion (uncued: '*Clever Kim called the cat clinic*', '*Sally sells seven spices*', '*Show Shelley the shady shoe shine*', '*Take today's tasty tea on the terrace*', '*The nightly news is never nice*', '*You used the yellow yoyo last year*'). The participants followed the instruction to read the sentences at their 'normal rate and loudness.' The goal of these recordings was to assess the stability of participants' articulatory performance before the beginning of the treatment.

At the third baseline session, four sentences were used to assess how participants responded to a verbal cue to 'use large speech movements' (cued: '*Clever Kim called the cat clinic*', '*Show Shelley the shady shoe shine*', '*Jimmy worked on a crossword puzzle*', '*That's my favourite Italian restaurant*'). The cued sentences were recorded in a blocked fashion, with four repetitions in the habitual style followed by four cued repetitions. The verbal cue was to 'use large speech movements'and was not supplemented with additional cues or feedback. Two of the cued sentences were subsequently trained during treatment (cued-trained: '*Jimmy worked on a crossword puzzle*', '*That's my favourite Italian restaurant*'), while two remained untrained (cued-untrained: '*Clever Kim called the cat clinic*', '*Show Shelley the shady show shine*').

The assessment stimuli (uncued, cued-trained, cued-untrained) were selected to represent a range of lingual consonants and both high and low vowels to allow participants recruit large articulatory movements. Since the focus of treatment was on increasing articulatory movement size as a whole, the primary criterion for stimuli selection was based on phonetic contexts that allowed for changes in movement size. The stimuli in this study were not otherwise controlled for linguistic or motoric complexity. All of the assessment stimuli were recorded in the absence of AVF.

Treatment. Treatment began immediately following the third baseline assessment. All sessions were conducted on an individual basis in a speech laboratory. The goal of treatment was to increase AWS of the tongue during sentence production when prompted with a verbal cue to ‘use large speech movement’.

Schedule and Stimuli. All participants attended 10 treatment sessions lasting approximately 45 minutes. Median treatment intensity was 1.5 sessions per week (*IQR* = 1.2-1.9). Throughout treatment, 50 functional sentences (five per session) were trained in random order (Appendix). Each sentence (including the two cued-trained sentences) was trained once over the course of treatment. The treatment stimuli were developed specifically for this study, and the range of sentences ensured that the ‘large movement’ strategy was practiced across phonetic contexts and sentence types (e.g., statements, questions). The sentences varied in length from six to 12 syllables (mean = 8.5, *SD* = 1.32), but were not controlled for linguistic or motoric complexity.

Protocol. Training of each sentence was conducted in three distinct phases: *calibration, acquisition* and *test* phases (Figure 2). The goal of the calibration phase was to set speaker and sentence specific AWS targets for training and to calibrate these targets within the AVF-game space. This phase was necessary due to anatomical differences between speakers, variability due to sensor placement, and phonetic differences between sentences. To establish the target AWS, participants produced the target sentence three times in their habitual style (uncued) and without AVF.

The acquisition phase aimed to train the participants to expand their articulatory movement relative to the calibration movement size and to sustain this expansion across repetitions. An initial target was specified as a 45% (±10%) increase in AWS from the median of the three calibration productions. Data from our previous pilot study indicated that a 45% increase was a reasonable target for the majority of speakers to attain (Yunusova et al., 2017). Participants were verbally cued to use large tongue movements, and terminal feedback regarding the target and achieved AWSs was provided following each sentence. As such, the feedback corresponded to both knowledge of results (above or below target) and knowledge of performance (magnitude of movement size). Terminal feedback was selected over a real-time display of feedback, as participants were required to read the treatment stimuli from the screen, and paying attention to feedback at the same time would have increased the attentional demands of the task (e.g., O'Shea, Morris, & Iansek, 2002).

Following the initial target set at a 45% increase in AWS, target setting depended on the participant's performance, and adapted based on the running mean of the previous three repetitions: 1) if the running mean was on target, the target level remained the same; 2) if the running mean exceeded the target, the target increased to the running mean (±10%); 3) if the running mean was less than the target, the target decreased by 15% (±10%). These reference values were empirically determined during our previous research (Yunusova et al., 2017). The acquisition phase was complete when participants successfully produced five consecutive repetitions of the target AWS. Alternatively, if performance did not stabilize by the 20th repetition, participants automatically progressed to the test phase. On average, participants required 9.64 repetitions ($SD = 3.35$; mean range across subjects = 8.1-12.6) to complete the acquisition phase per sentence and rarely (8/250 sentences) reached the 20th repetition without stabilizing their AWS.

The goal of the test phase was to encourage participants to use the verbal cue without depending on visual feedback. The final target (AWS) setting, obtained in the acquisition phase, was carried forward to the test phase. Participants were cued to use large tongue movements for six repetitions, and feedback was provided on a reduced schedule (50% of trials). Participants selected the three trials to receive visual feedback on, in order to increase motivation and engagement with the learning process (Chiviacowsky, Wulf, Lewthwaite, & Campos, 2012). Performance on the trials that followed the selection of feedback (or no feedback) allowed an evaluation of the effect of AVF on tongue movement size during treatment. Summary feedback regarding performance during the test phase was shown at the end of the phase (e.g., 4/6).

Visual Feedback During Treatment. Two games that were developed in-house (Haworth, 2016) – one representing a "dragon world" and one a "fish world" – were used on alternating treatment days (Figure 3). In the dragon world, the extent of fire breathed by a dragon, corresponding to the AWS of the tongue, was shown, as well as the location of a target object to burn (e.g., a wooden barrel). Similarly, in the fish world, a fishing net corresponded to the size of the tongue AWS, and the target was indicated by different types of fish. The expansion of the fire in the "dragon world" occurred primarily in the horizontal dimension, whereas the net expansion in the "fish world" occurred in both the horizontal and vertical dimensions. In both worlds, the visual game space allowed for an increase in movement size of approximately 450% from the participant's baseline. Above this threshold, the extent of the fire and fishing net went offscreen. Each world had five levels, and participants progressed from one level to the next after each session. At the end of each session, a cumulative score from all test phases was shown on a "high score board." Participants were able to see the scores from all players (anonymized) on the scoreboard.

Post-Treatment Assessment. The assessment procedures following treatment were identical to the third baseline session, in order to evaluate the effects of treatment. The six uncued sentences were recorded in the habitual style, and the four cued sentences were recorded, first in the habitual style, and then with a cue to 'use large speech movements'.

Speech Intelligibility Ratings Pre and Post Treatment

Five naïve listeners (F = 5; mean age = 27.72 ± 4.16) were recruited to rate intelligibility before and after treatment. Given the subjective nature of perceptual ratings, multiple listeners were required to provide an overall estimate of speech intelligibility. All listeners pases a pure-tone hearing screen at 20dB HL for frequencies ranging from 250-8000Hz bilaterally. The listeners were native speakers of English, had at least a high school diploma, and reported no history of speech or language disorders.

The audio recordings were post-processed before intelligibility rating using Goldwave Version 6 software (Goldwave Inc, 2015); non-speech high-frequency noise attributed to the WAVE was removed from the signal (low-pass filter at 9800Hz), and the recordings were equated for root mean square amplitude to minimize intelligibility effects due to audibility (Tjaden, Sussman, & Wilding, 2014 2014). The stimuli were then mixed with speech-shaped noise at a signal-to-noise-ratio (SNR) of -5dB (van Engen, Phelps, Smiljanic, & Chandrasekaran, 2014), in order to avoid a ceiling effect in the data and to create a listening environment that more closely resembles everyday communication situations.

The listeners were asked to perform paired-comparison ratings of intelligibility of the two cued-trained sentences from pre and post-treatment sessions (Park, Theodoros, Finch, & Cardell, 2016; Wenke, Theodoros, & Cornwell, 2011). The purpose of this task was to assess if increases in tongue movement size corresponded to improvements in speech intelligibility following treatment. As such, the post-treatment samples were selected based on a minimum increase of 1*SD* in tongue movement size from the pre-treatment mean.

One repetition of each sentence pre and post treatment was randomly selected for each speaker. Each pair of speech samples was presented to listeners in both pre-post and post-pre combinations in random order. Each pair, therefore, was rated twice by each listener. The listeners were required to decide whether the first or second sample of each pair was easier to understand, or if there was no perceptible difference between them (they sounded the same). The task instructions were adapted from previous studies that used paired-comparison ratings of intelligibility (Park et al., 2016; Wenke et al., 2011). The listeners were blinded to the assessment time of the recordings (i.e., pre vs. post). The recordings were presented once through headphones (BOSE QuietComfort 15) in a sound-attenuated booth (Industrial Acoustics Co.) using E-Prime 2.0 experiment software (Psychology Software Tools Inc, 2012). Before completing the experimental task, the listeners practiced rating five pairs of audio recordings that were not part of the current study to ensure that they understood the requirements of the task. They were allowed to ask questions to clarify the task instructions but were not given specific feedback regarding their ratings. As each pair of pre-post speech samples was rated twice by each listener, a total of 100 ratings were obtained (5 speakers x 2 sentences x 2 presentation orders x 5 listeners).

All ratings were assessed for intra-rater reliability using Cohen's kappa and inter-rater reliability using Fleiss' kappa for multiple raters. The reliability coefficients were interpreted using benchmarks proposed by Landis and Koch (1977). The intra-rater reliability analysis revealed kappa (*k*) coefficients ranging from .62 to .77 (mean = .68, *SD* = .06), indicating substantial agreement within listeners. The inter-rater reliability analysis showed a k value of .22, indicating a fair agreement between listeners.

Data Analysis

Baseline Performance. Visual analysis of AWS in the uncued sentences was conducted across the three baseline sessions to informally assess the stability of articulatory performance before treatment, i.e., to ensure that participants were not increasing their AWS before starting treatment. AWS data were mean-centered for this analysis to account for the inherent differences in movement size between sentences, and the mean and standard deviations across sentences were compared graphically across the baseline sessions. Additionally, effect sizes for the cued sentences (uncued-cued productions) were calculated to evaluate participants' response to the verbal cue alone before treatment. The magnitude of effect was determined using a variation of Cohen's *d* statistic, which pools an individual's standard deviation across conditions (d_2) (Busk & Serlin, 1992). Currently, there is no empirically established bench-mark for interpreting effect sizes for AWS data. In the limit, effect sizes greater than 1 were interpreted as a clinically significant difference. An effect size greater than 1 indicates that the difference between mean values exceeds the pooled standard deviation (Maas & Farinella, 2012).

Effect of Augmented Visual Feedback During Treatment. Data from the test phase (after acquisition) of the treatment sessions were visually analyzed to assess the effect of AVF on articulatory movements during treatment. Specifically, articulatory performance on the trials that followed feedback was compared to the trials that did not follow feedback. Percent change in

AWS was calculated for the feedback and no feedback trials from the calibration (uncued) to test (cued) phase of each sentence. Average percent change values for the feedback and no feedback trials were then computed per session and plotted as time series. The visual analysis was supplemented with the two-standard deviation (2SD) band analysis method (Bloom, 1975; Nourbakhsh & Ottenbacher, 1994). Using this method, a line representing 2*SD* above the pre-treatment mean was added to the time series; two consecutive treatment points above this line indicates a significant improvement in performance.

Pre-Post Treatment Effect. The magnitude of treatment and generalization effects for AWS data pre-post treatment was determine using effect sizes, as described above. Three effects sizes of interest were calculated pre-post treatment for each participant: (1) cued-trained sentences to assess treatment effect; (2) cued-untrained sentences to assess generalization from trained to untrained sentences; and (3) uncued sentences to assess generalization to habitual speech.

Paired-comparison ratings of speech intelligibility for the cued-trained sentences were examined descriptively. All statistical and graphical analyses were conducted in R version 3.3.2 (R Core Team, 2016).

Results

Articulatory Working Space

Baseline Performance. Figure 5 shows mean-centered AWS data for the three baseline sessions for each participant in their habitual style (uncued). Three participants (PD14, PD25, PD30) were judged to demonstrate a relatively stable AWS measure across sentences before treatment. Two participants (PD27, PD28) showed variable performance, but visual inspection did not suggest either a rising or falling trend for either participant.

Figure 6 shows percent change in AWS for each participant in response to a verbal cue at baseline (uncued-cued productions of cued sentences). Table 2 shows the corresponding effect sizes. One participant (PD25) showed a large and clinically significant effect size in response to the verbal cue alone before treatment. Two participants (PD28, PD30) showed small effect sizes, and two participants (PD14, PD27) had difficulty applying the large movement instruction to their articulatory performance and were unable to increase their AWS.

Effect of Augmented Visual Feedback During Treatment. Figure 7 shows average percent change in AWS for sentences produced during treatment, with and without visual feedback. Four of five participants showed improved performance during the trials that followed feedback by increasing their AWS to a greater extent than their pre-treatment (cued) levels for at least two consecutive sessions. The same participants also exceeded this threshold for the majority of the sessions (PD14, 6/10 sessions; PD25, 8/10 sessions; PD27, 9/10 sessions; PD30, 10/10 sessions). In comparison, only one participant improved their performance during the trials following no feedback, performing above the threshold for 7/10 sessions (PD27).

The extent of increase in AWS varied considerably across participants, particularly in the feedback condition (Table 3). On average, PD28 showed the smallest percent increase in AWS (Feedback, 46%; No feedback, 27%) and PD30 showed the largest percent increase in AWS (Feedback, 770%; No feedback, 83%).

Pre-Post Treatment Effect. Effect sizes representing a change in AWS from pre to post treatment are shown for all participants in Table 4. The majority of participants responded to treatment, showing large and clinically significant effect sizes for cued-trained sentences (n = 4). Two of five participants generalized the effect of treatment to cued-untrained stimuli. No evidence of generalization to uncued sentences was observed for any participant after treatment.

Speech Intelligibility

Table 5 shows the percentage of paired-comparison ratings for each participant that were judged as easier to understand pre or post treatment, or as being the same pre-post treatment, for CUED-trained sentences. One participant (PD28) was rated as being more intelligible post treatment and one was rated as the same pre-post treatment (PD27). The other three participants were rated as being less intelligible after the treatment (PD14, PD25, PD30).

Discussion

The present study investigated the effects of using a verbal cue in conjunction with AVF to treat articulatory hypokinesia in five individuals with PD. The premise of this study was that hypokinesia is a contributing factor to the speech impairment exhibited by the individuals with PD (Kearney et al., 2017). We set to evaluate whether and how articulatory movement size can be changed in a therapeutic context, and whether this change would have a positive effect on speech intelligibility. At baseline, participant responses to the verbal cue alone to increase speech movements were relatively limited. The training with AVF resulted in a substantial increase in AWS in 4/5 participants beyond their baseline response to a verbal cue. Two of these participants successfully generalized the ability to use the cue to untrained sentences. None of the participants, however, applied the strategy to their habitual style of speech, when the verbal cue was removed. Changes in speech intelligibility examined in cued sentences pre and post treatment were not in an expected direction – in three participants, an increase in AWS was associated with worsening of speech intelligibility. These findings are discussed below with respect to the future design of speech therapy in PD and theoretical implications regarding articulatory targets for speech rehabilitation.

The Role of Cueing and AVF in Training Larger Speech Movements

The limited response to a verbal cue to increase speech movements at baseline was observed in this study. This finding is in contrast to previous findings that showed that speakers with PD could increase their articulatory movements when cued to speak louder or more clearly, albeit to a lesser extent than control speakers (Darling & Huber, 2011; Dromey, 2000; Kearney et al., 2017). These cues, however, vary in the nature of the target with the current cue focusing on somatosensory information in comparison to the auditory focus of speaking louder or, possibly, more clearly. While individuals may have experience in their daily lives in adjusting their speech to effect the acoustic signal (e.g., speaking louder in a noisy environment), it is likely that changing the size of their articulatory movements is a novel task and one that does not have an obvious reference or target point. Further, the task may be difficult for patients with PD who experience deficits in proprioception and sensorimotor integration required for speech (Mollaei et al., 2013; Schneider et al., 1986).Because a verbal cue alone had a limited effect on articulatory kinematicsat baseline, and monitoring of spatial features is difficult in a clinical setting without the help of instrumentation (Simione et al., 2016), an instrumentation-based AVF treatment was employed.

During treatment, AVF was more effective than a verbal cue alone in increasing AWS (average increase with AVF: 270.7%; without AVF: 51.9%). Following treatment, participants were also able to use larger tongue movements – when cued – than before treatment. These findings were anticipated given the reported success of AVF in enhancing motor learning and treatment outcomes for individuals with PD (see reviews, Kearney et al., 2018; Nieuwboer, Rochester, Muncks, & Swinnen, 2009). AVF may have been particularly important in the current treatment approach because conceptualizing AWS is not a typical process in normal speech production. AVF regarding articulatory movement size seemed to facilitate this conceptualization by providing a reference for performance over the course of the treatment. In subsequent trials, the participants were able to apply a corrective response in planning their next movements, which may have helped to strengthen the feedforward control of their movement, and to use it when AVF was no longer available (Perkell, 2012).

The Effect of Larger Speech Movements on Speech Intelligibility

Although AVF appeared to be beneficial in teaching participants to increase their articulatory movements, this increase had varied effects on speech intelligibility across speakers. Only one participant showed improved speech intelligibility after treatment and three showed a detriment to their speech intelligibility.

The participant with the smallest increase in AWS (PD28) was the only one showing improvement in speech intelligibility post treatment. His average increase in AWS was 36.9% (*SD* = 15.7%) over the course of the treatment. It is important to remember that in the present design, there was no upper limit set for the enlargement of AWS, which meant that during treatment, participants were increasing their movement size without a pre-defined maximum target; as such, they ended up increasing their AWS by an average of 161.3%. This increase in movement size may reflect the maximum they could achieve within their anatomical constraints. The focus on speech movement size may have reduced the speakers' attention on the resulting acoustic signal, particularly when self-monitoring of the acoustic signal may be further impaired in patients with PD, who are reported to experience deficits in auditory processing (Kwan & Whitehill, 2011). These data suggested that above a certain point, increased movement amplitude might result in speech sounding less natural and more difficult to understand and, therefore, movement size boundaries may need to be set and monitored carefully in a movement-based treatment.

The suggestion that an optimal range of articulatory performance may exist is theoretically rooted in the DIVA model of speech production (Guenther, 1995). Guenther (1995) proposed that there exist target ranges, along which articulatory positions can vary while still producing the same target. In addition to established target ranges for the production of specific vowels and consonants (Perkell, 1996; Yunusova, Rosenthal, Rudy, Baljko, & Daskalogiannakis, 2012), the current data also suggest that a target range may exist for articulatory movement size at the sentence level. Articulatory movements above or below this optimal range appear to be associated with reduced intelligibility. Given the non-linear relationship observed between kinematics and intelligibility, the articulatory-expansion treatment should establish a method for identifying the optimal articulatory movement range. Examining within-person changes in speech movement size with respect to changes in their speech intelligibility may be a plausible way of identifying target ranges for therapy.

In the past, the targets of speech production have been classified as either acoustic or somatosensory in nature, and the primary nature of either has been debated in a number of studies (Perkell et al., 1997; Saltzman & Munhall, 1989), with a current consensus on the importance of both (Houde & Jordan, 1998; Nasir & Ostry, 2008; Tourville, Reilly, & Guenther, 2008; Tremblay, Shiller, & Ostry, 2003). Little attention, however, has been given to the notion of speech targets in the context of speech therapy. The speech intelligibility results of this study emphasize the importance of focusing on articulatory kinematics in order to achieve movement changes but linking the articulatory and acoustic targets to optimize the effect of articulatory changes on speech quality.

Generalization of Treatment Effects to Untrained and Uncued Contexts

The goal of rehabilitation is to promote generalization of treatment gains to untrained stimuli and other contexts (Ballard, 2001). In the current study, two participants were able to generalize the treatment effect to untrained stimuli, but none generalized the large movement strategy to their habitual speech. It is important to consider here the distinction between the acquisition of motor skills, as shown in the treatment effect data, and the generalization of those skills to other stimuli/contexts. While motor skill acquisition can result in a change in performance, evidence of transfer or generalization is required to demonstrate that motor learning has occurred, and reflects relatively permanent changes in the general capability for movement (Schmidt & Lee, 2011). Our generalization findings may indicate a need to modify aspects of the treatment design to further enhance motor learning.

A potential factor to consider is the treatment schedule and whether 10 sessions (45 mins each) over five weeks was a sufficient amount and intensity of practice to facilitate generalization of treatment effects. Previous studies have indicated that individuals with PD experience significant difficulties at the automatization stage of learning and require extended practice to achieve this level of motor skill (Nieuwboer et al., 2009). Treatments targeting loud speech in individuals with PD have established a treatment schedule of 16 sessions over four weeks (Martens et al., 2015; Ramig et al., 1995), which may be required for generalization of effects to occur. Such a schedule would be facilitated by home-based practice. The high hardware and software requirements for the current treatment set-up would make it difficult to implement this treatment in a clinic or home setting. Our group, however, is currently investigating the validity of alternative technologies, such as facial tracking, that could be used in the future implementation of this treatment (Bandini, Namasivayam, & Yunusova, 2017).

Further, motor learning is considered to be specific to the conditions of training (specificity of practice; Proteau, 1992), and this principle of motor learning may explain why none of the participants habituated the large movement strategy to uncued sentences following treatment. During treatment, the participants were always cued to use large articulatory movements. As a result, the effect of treatment was observed for cued sentences post treatment, but not for the uncued sentences. Refining the protocol to include practice opportunities without a verbal cue may encourage participants to apply the large movement strategy to their habitual speaking style.

Limitations and Future Directions

In addition to the lack of the upper limit target during treatment and factors related to treatment design (e.g., the frequency of cueing), a few further limitations might need to be considered when interpreting the results and designing future treatment studies in this population. First, the assessment of performance during treatment was conducted across different stimuli in every session. To compare across sessions, we examined percent change in movement size from uncued to cued productions; however, we were unable to account for phonetic differences across sentences that may have facilitated (or hindered) AWS expansion. As a result, it is difficult to explain some of the variability observed across treatment sessions. A future study of this treatment will incorporate a systematically-designed probe list with the same items to be administered at regular intervals throughout treatment.

The intelligibility findings of this study stress the importance of assessing speech intelligibility in motor speech treatment research, which is not commonly implemented. Measuring intelligibility is not a trivial task (Kent, 1996; Miller, 2013). We employed untrained, naive listeners that attempted a global judgement through comparisons of sentence pairs (Park et al., 2016; Wenke et al., 2011); however, the modest inter-rater reliability results may suggest that a clear distinction was not present between pre and post-treatment recordings. Some of the listeners commented that the recordings sometimes varied in rate/intonation but that these differences did not make a recording easier to understand. Transcription-based methods by highly trained listeners may show greater sensitivity to sound-specific changes that occur in articulation post treatment (Miller, 2013). Establishing reliable methods of assessing within-subject change due to treatment is pertinent to understanding the effect of treatment on speech intelligibility, and linking changes in intelligibility to underlying changes in articulatory movements. Further, the inclusion of other clinically relevant measures, such as articulatory rate, may shed light on treatment effects not captured by ratings of speech intelligibility.

Given the small sample size in the current study, the results provide only a preliminary level of evidence regarding the effect of AVF on tongue AWS and speech intelligibility in PD, and the generalizability of findings is limited. In addition, variability in treatment responsiveness was observed but cannot be fully explained in this very small group of participants. Future studies examining physiological changes in articulation pre-post treatment will require a larger group of participants to generalize findings and delineate factors associated with treatment candidacy.

Although the descision to use AWS as a treatment target is rooted in the current understanding of the speech movement disorder underlying dysarthria in PD, other targets may be equally justified. For example, improved control of movement velocity/speed may also be a reasonable target, given that tongue movement speed has previously been shown to be correlated to speech intelligibility across speakers with PD (Weismer et al., 2012).

Finally, the current study did not include a comparison to a traditional treatment used in this population, such as LSVT LOUD, or to a similar treatment based on articulary expansion without AVF (e.g., LSVT-ARTIC, for which published empirical data are not yet available). As a result, it is not possible to determine whether the direct amplitude training approach with AVF is more or less effective in addressing articulatory hypokinesia in PD than other pretment approaches. In the long term, it will be important to conduct efficacy trials comparing the novel treatment approach to traditional interventions in order to further inform evidence-based practice. Careful descriptions of speech disorder presentation as well as participant selection might be required to determine treatment candidacy for different intervention appraoches.

Conclusion

Treatment approaches that directly address the articulatory impairment in individuals with PD remain limited. The present study is an initial step in the programmatic evaluation of a movement-based intervention using AVF for speech rehabilitation in individuals with PD, who exhibit speech intelligibility impairment due to hypokinesia. AVF may be beneficial in speech therapy in order to train an increase in articulatory movement size – assuming boundary conditions are empirically defined – and to maintain articulatory function. The focus on the quality of the acoustic signal, however, may also be imperative. Further modifications to target specification are required to optimize the effects of treatment on speech intelligibility, and novel technological solutions to acquiring speech movement data are needed to improve the feasibility of treatment delivery.

Acknowledgements

Portions of this study were presented at the 7th International Conference on Speech Motor Control in Groningen, The Netherlands in July 2017.This research was supported by the

Parkinson's Society of Canada Pilot Project Grant, the Natural Sciences and Engineering Research Council (NSERC) Discovery Grant, and the Centre for Innovation in Information Visualization and Data-Driven Design (CIV-DDD). The first author was also supported by a National Institute on Deafness and Other Communication Disorders Grant (R01 DC016270), awarded to C. Stepp and F. Guenther. We are grateful to the participants and their families for taking part in this study. We also thank Madhura Kulkarni and Vincci Tau for their assistance with this project.

References

Ackermann, H., Konczak, J., & Hertrich, I. (1997). The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and Language, 56*(2), 312-319. doi:10.1006/brln.1997.1851

Atkinson-Clement, C., Sadat, J., & Pinto, S. (2015). Behavioral treatments for speech in Parkinson's disease: meta-analyses and review of the literature. *Neurodegenerative Disease Management, 5*(3), 233-248.

Ballard, K. J. (2001). Response generalization in apraxia of speech treatments: taking another look. *Journal of Communication Disorders, 34*(1), 3-20. doi:10.1016/S0021-9924(00)00038-1

Bandini, A., Namasivayam, A., & Yunusova, Y. (2017, August). *Video-based tracking of jaw movements during speech: Preliminary results and future directions.* Paper presented at the INTERSPEECH 2017 - 18th Annual Conference of the International Speech Communication Association, Stockholm, Sweden.

Beall, E. B., Lowe, M. J., Alberts, J. L., Frankemolle, A. M., Thota, A. K., Shah, C., & Phillips, M. D. (2013). The effect of forced-exercise therapy for Parkinson's disease on motor cortex functional connectivity. *Brain connectivity, 3*(2), 190-198.

Berry, J. J. (2011). Accuracy of the NDI wave speech research system. *Journal of Speech, Language, and Hearing Research, 54*(5), 1295-1301. doi:10.1044/1092-4388(2011/10-0226)

Bertram, L., & Tanzi, R. E. (2005). The genetic epidemiology of neurodegenerative disease. *The Journal of Clinical Investigation, 115*(6), 1449-1457. doi:10.1172/JCI24761

Bloom, M. (1975). *The paradox of helping: Introduction to the philosophy of scientific practice*. New York, NY: Wiley.

Busk, P. L., & Serlin, R. (1992). Meta-analysis for single-case research. In T. Kratochwill & J. Levin (Eds.), *Single case research design and analysis* (pp. 187-212). Hillsdale, NJ: Erlbaum.

Campbell, T. F., & Yunusova, Y. (2017). Cutting-edge technology for speech rehabilitation. *Journal of Speech, Language, and Hearing Research, 60*(6S), 1798-1799.

Chiviacowsky, S., Wulf, G., Lewthwaite, R., & Campos, T. (2012). Motor learning benefits of self-controlled practice in persons with Parkinson's disease. *Gait & Posture, 35*(4), 601-605. doi:10.1016/j.gaitpost.2011.12.003

Cleland, J., Scobbie, J. M., & Wrench, A. A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics, 29*(8-10), 575-597.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech, Language, and Hearing Research, 12*(3), 462-496. doi:10.1044/jshr.1203.462

Darling, M., & Huber, J. E. (2011). Changes to articulatory kinematics in response to loudness cues in individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 54*(5), 1247-1259. doi:10.1044/1092-4388(2011/10-0024)

de Swart, B. J., Willemse, S., Maassen, B., & Horstink, M. (2003). Improvement of voicing in patients with Parkinson's disease by speech therapy. *Neurology, 60*(3), 498-500. doi:10. 1212/01.WNL.0000044480.95458.56

Dent, H., Gibbon, F., & Hardcastle, B. (1995). The application of electropalatography (EPG) to the remediation of speech disorders in school-aged children and young adults. *International Journal of Language & Communication Disorders, 30*(2), 264-277.

Desmurget, M., Grafton, S., Vindras, P., Grea, H., & Turner, R. S. (2004). The basal ganglia network mediates the planning of movement amplitude. *European Journal of Neuroscience, 19*(10), 2871-2880.

Dromey, C. (2000). Articulatory kinematics in patients with Parkinson disease using different speech treatment approaches. *Journal of Medical Speech-Language Pathology, 8*, 155-161.

Ebersbach, G., Ebersbach, A., Edler, D., Kaufhold, O., Kusch, M., Kupsch, A., & Wissel, J. (2010). Comparing exercise in Parkinson's disease—the Berlin BIG Study. *Movement Disorders, 25*(12), 1902-1908.

Farley, B. G., Fox, C. M., Ramig, L. O., & McFarland, D. H. (2008). Intensive Amplitude-specific Therapeutic Approaches for Parkinson's Disease: Toward a Neuroplasticity-principled Rehabilitation Model. *Topics in Geriatric Rehabilitation, 24*(2), 99-114.

Ford, M. P., Malone, L. A., Nyikos, I., Yelisetty, R., & Bickel, C. S. (2010). Gait training with progressive external auditory cueing in persons with Parkinson's disease. *Archives of Physical Medicine and Rehabilitation, 91*(8), 1255-1261.

Forrest, K., Weismer, G., & Turner, G. S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults. *Journal of Acoustical Society of America, 85*(6), 2608-2622. doi:10.1121/1.397755

Fox, C., Ebersbach, G., Ramig, L. O., & Sapir, S. (2012). LSVT LOUD and LSVT BIG: Behavioural treatment programs for speech and body movement in Parkinson disease. *Parkinson's Disease, 2012*, 1-12. doi:10.1155/2012/391946

Goldwave Inc. (2015) (Version 6) [Computer software]. St. John's, Canada: Goldwave Inc. Retrieved from http://www.goldwave.ca/release.php

Gracco, V. L. (1992). Analysis of Speech Movements: Practical Considerations and Clinical Application. *Haskins Laboratories Status Report on Speech Research, SR-109/110*, 45-58.

Grimme, B., Fuchs, S., Perrier, P., & Schöner, G. (2011). Limb versus speech motor control: A conceptual review. *Motor Control, 15*(1), 5-33.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*(3), 594-621. doi:10.1037/0033-295X.102.3.594

Haworth, B. (2016). *Computer games for motor speech rehabiliation* (Master's thesis). York University, Toronto, Canada. Retrieved from http://hdl.handle.net/10315/32241

Haworth, B., Kearney, E., Baljko, M., Faloutsos, P., & Yunusova, Y. (2014). *Electromagnetic articulography in the development of 'serious games' for speech rehabiliation.* Paper presented at the The 2nd International Workshop on BioMechanical and Parametric Modelling of Human Anatomy, Vancouver, Canada.

Haworth, B., Kearney, E., Faloutsos, P., Baljko, M., & Yunusova, Y. (2018). Electromagnetic articulography (EMA) for real-time feedback application: Computational techniques. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*. doi:10.1080/21681163.2018.1434423

Ho, A. K., Iansek, R., Marigliani, C., Bradshaw, J. L., & Gates, S. (1998). Speech impairment in a large sample of patients with Parkinson's disease. *Behavioural Neurology, 11*(3), 131-137. doi:10.1155/1999/327643

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279*(5354), 1213-1216. doi:10.1126/science.279.5354.1213

Karlsen, K., Larsen, J. P., Tandberg, E., & Mæland, J. G. (1999). Influence of clinical and demographic variables on quality of life in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 66*(4), 431-435.

Katz, W. F., Bharadwaj, S. V., & Carstens, B. (1999). Electromagnetic articulography treatment for an adult with Broca's aphasia and apraxia of speech. *Journal of Speech, Language, and Hearing Research, 42*(6), 1355-1366.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.

Kearney, E., Giles, R., Haworth, B., Faloutsos, P., Baljko, M., & Yunusova, Y. (2017). Sentence-level movements in Parkinson's disease: Loud, clear, and slow speech. *Journal of Speech, Language and Hearing Research, 60*, 3426-3440. doi:10.1044/2017_JSLHR-S-17-0075

Kearney, E., Shellikeri, S., Martino, R., & Yunusova, Y. (2018). Augmented visual feedback-aided interventions in motor rehabilitation. *Disability and Rehabilitation*. doi:10.1080/09638288.2017.1419292

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology, 5*(3), 7-23.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. doi:10.2307/2529310

Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *Journal of Speech and Hearing Disorders, 43*(1), 47-57. doi:10.1044/jshd.4301.47

Lowit, A., Dobinson, C., Timmins, C., Howell, P., & Kröger, B. (2010). The effectiveness of traditional methods and altered auditory feedback in improving speech rate and intelligibility in speakers with Parkinson's disease. *International Journal of Speech-Language Pathology, 12*(5), 426-436. doi:10.3109/17549507.2010.497559

Maas, E., & Farinella, K. A. (2012). Random versus blocked practice in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55*(2), 561-578. doi:10.1044/1092-4388(2011/11-0120)

Martens, H., Van Nuffelen, G., Dekens, T., Huici, M. H.-D., Hernández-Díaz, H. A. K., De Letter, M., & De Bodt, M. (2015). The effect of intensive speech rate and intonation therapy on intelligibility in Parkinson's disease. *Journal of Communication Disorders, 58*, 91-105. doi:10.1016/j.jcomdis.2015.10.004

McNeil, M., Katz, W., Fossett, T., Garst, D., Szuminsky, N., Carter, G., & Lim, K. (2010). Effects of online augmented kinematic and perceptual feedback on treatment of speech movements in apraxia of speech. *Folia Phoniatrica et Logopaedica, 62*(3), 127-133.

McRae, P. A., Tjaden, K., & Schoonings, B. (2002). Acoustic and perceptual consequences of articulatory rate change in Parkinson disease. *Journal of Speech, Language, and Hearing Research, 45*(1), 35-50. 2012/08/24/15:00:15. doi:10.1044/1092-4388(2002/003)

Miller, N. (2013). Measuring up to speech intelligibility. *International Journal of Language & Communication Disorders, 48*(6), 601-612. doi:10.1111/1460-6984.12061

Mollaei, F., Shiller, D. M., & Gracco, V. L. (2013). Sensorimotor adaptation of speech in Parkinson's disease. *Movement Disorders, 28*(12), 1668-1674. doi:10.1002/mds.25588

Nackaerts, E., Nieuwboer, A., & Farella, E. (2017). Technology-Assisted Rehabilitation of Writing Skills in Parkinson's Disease: Visual Cueing versus Intelligent Feedback. *Parkinson's Disease, 2017*.

Nasir, S. M., & Ostry, D. J. (2008). Speech motor learning in profoundly deaf adults. *Nature Neuroscience, 11*(10), 1217-1222. doi:10.1038/nn.2193

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699. doi:10.1111/j.1532-5415.2005.53221.x

Nieuwboer, A., Rochester, L., Muncks, L., & Swinnen, S. P. (2009). Motor learning in Parkinson's disease: limitations and potential for rehabilitation. *Parkinsonism & Related Disorders, 15*(s3), s53-58. doi:10.1016/S1353-8020(09)70781-3

Nourbakhsh, M. R., & Ottenbacher, K. J. (1994). The statistical analysis of single-subject data: a comparative examination. *Physical Therapy, 74*(8), 768-776. doi:10.1093/ptj/74.8.768

O'Shea, S., Morris, M. E., & Iansek, R. (2002). Dual task interference during gait in people with Parkinson disease: effects of motor versus cognitive secondary tasks. *Physical Therapy, 82*(9), 888-897. doi:10.1093/ptj/82.9.888

Oliveira, R. M., Gurd, J. M., Nixon, P., Marshall, J. C., & Passingham, R. E. (1997). Micrographia in Parkinson's disease: the effect of providing external cues. *Journal of Neurology, Neurosurgery & Psychiatry, 63*(4), 429-433.

Park, S., Theodoros, D., Finch, E., & Cardell, E. (2016). Be clear: A new intensive speech treatment for adults with nonprogressive dysarthria. *American Journal of Speech-Language Pathology, 25*(1), 97-110. doi:10.1044/2015_AJSLP-14-0113

Perkell, J. (1996). Properties of the tongue help to define vowel categories: hypotheses based on physiologically-oriented modeling. *Journal of Phonetics, 24*(1), 3-22. doi:10.1006/jpho.1996.0002

Perkell, J. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics, 25*(5), 382-407. doi:10.1016/j.jneuroling.2010.02.011

Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22*(2), 227-250. doi:10.1016/S0167-6393(97)00026-5

Proteau, L. (1992). On the specificity of learning and the role of visual information for movement control. *Advances in Psychology, 85*, 67-103. doi:10.1016/S0166-4115(08)62011-7

Psychology Software Tools Inc. (2012). E-Prime (Version 2.0) [Computer software]. Sharpsburg, PA: Psychology Software Tools, Inc. Retrieved from http://www.pstnet.com/eprime.cfm

R Core Team. (2016). R: A language and environment for statistical computing. (Version 3.3.2) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

Ramig, L. O., Countryman, S., Thompson, L. L., & Horii, Y. (1995). Comparison of two forms of intensive speech treatment for Parkinson's disease. *Journal of Speech and Hearing Research, 38*(6), 1232-1251. doi:10.1044/jshr.3806.1232

Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A., O'brien, C., Hoehn, M., & Thompson, L. (2001). Intensive voice treatment (LSVT®) for patients with Parkinson's disease: a 2 year follow up. *Journal of Neurology, Neurosurgery & Psychiatry, 71*(4), 493-498. doi:10.1136/jnnp.71.4.493

Richardson, K., Sussman, J. E., Stathopoulos, E. T., & Huber, J. E. (2014). The effect of increased vocal intensity on interarticulator timing in speakers with Parkinson's disease: A preliminary analysis. *Journal of Communication Disorders, 52*, 44-64. doi:10.1016/j.jcomdis.2014.09.004

Robey, R. R. (2004). A five-phase model for clinical-outcome research. *Journal of Communication Disorders, 37*(5), 401-411. doi:10.1016/j.jcomdis.2004.04.003

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1*(4), 333-382. doi:10.1207/s15326969eco0104_2

Schmidt, R. A., & Lee, T. D. (2011). *Motor control and learning: A behavioral emphasis* (5 ed.). Champaign, IL: Human Kinetics

Schneider, J., Diamond, S. G., & Markham, C. H. (1986). Deficits in orofacial sensorimotor function in Parkinson's disease. *Annals of Neurology, 19*, 275-282. doi:10.1002/ana.410190309

Scott, S., & Caird, F. (1984). The response of the apparent receptive speech disorder of Parkinson's disease to speech therapy. *Journal of Neurology, Neurosurgery & Psychiatry, 47*(3), 302-304. doi:10.1136/jnnp.48.6.606

Shtern, M., Haworth, M., Yunusova, Y., Baljko, M., & Faloutsos, P. (2012). A game system for speech rehabilitation. In M. Kallmann & K. Bekris (Eds.), *Lecture Notes in Computer Science* (Vol. 7660, pp. 43-54). Berlin, Germany: Springer.

Simione, M., Wilson, E. M., Yunusova, Y., & Green, J. R. (2016). Validation of Clinical Observations of Mastication in Persons with ALS. *Dysphagia, 31*(3), 367-375. doi:10.1007/s00455-015-9685-3

Spaulding, S. J., Barber, B., Colby, M., Cormack, B., Mick, T., & Jenkins, M. E. (2013). Cueing and gait improvement among people with Parkinson's disease: a meta-analysis. *Archives of Physical Medicine and Rehabilitation, 94*(3), 562-570. doi:10.1016/j.apmr.2012.10.026

Swinnen, S. P. (1996). Information feedback for motor skill learning: A review. In H. N. Zelaznik (Ed.), *Advances in Motor Learning and Control* (pp. 37-66). Champaign, IL: Human Kinetics.

Tjaden, K., Sussman, J. E., & Wilding, G. E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research, 57*(3), 779-792. doi:10.1044/2014_JSLHR-S-12-0372

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage, 39*(3), 1429-1443. doi:10.1016/j.neuroimage.2007.09.054

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature, 423*(6942), 866-869. doi:10.1038/nature01710

Unity Technologies Inc. (2015). Unity 3D (Version 4.6.5p1) [Computer software]. San Francisco, CA: Unity Technologies Inc. Retrieved from https://docs.unity3d.com/460/Documentation/Manual/index.html

van Engen, K. J., Phelps, J. E., Smiljanic, R., & Chandrasekaran, B. (2014). Enhancing speech intelligibility: interactions among context, modality, speech style, and masker. *Journal of Speech, Language, and Hearing Research, 57*(5), 1908-1918. doi:10.1044/JSLHR-H-13-0076

Ventry, I., & Weinstein, B. (1983). Identification of elderly people with hearing problems. *American Speech-Language-Hearing Association, 25*(7), 37-42.

Walsh, B., & Smith, A. (2012). Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease. *Movement Disorders, 27*(7), 843-850. doi:10.1002/mds.24888

Weismer, G., & Laures, J. S. (2002). Direct magnitude estimates of speech intelligibility in dysarthria: Effects of a chosen standard. *Journal of Speech, Language and Hearing Research, 45*(3), 421-433. doi:10.1044/1092-4388(2002/033)

Weismer, G., Yunusova, Y., & Bunton, K. (2012). Measures to Evaluate the Effects of DBS on Speech Production. *Journal of Neurolinguistics, 25*(4), 74-94. doi:10.1016/j.jneuroling.2011.08.006

Wenke, R. J., Theodoros, D., & Cornwell, P. (2011). A comparison of the effects of the Lee Silverman voice treatment and traditional therapy on intelligibility, perceptual speech features, and everyday communication in nonprogressive dysarthria. *Journal of Medical Speech-Language Pathology, 19*(4), 1-25.

Yorkston, K. M., Beukelman, D., Hakel, M., & Dorsey, M. (2007). Sentence Intelligibility Test. Lincoln, NE.

Yorkston, K. M., Hakel, M., Beukelman, D., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology, 15*(2), xi-xxxvi.

Yunusova, Y., Kearney, E., Kulkarni, M., Haworth, B., Baljko, M., & Faloutsos, P. (2017). Game-based augmented visual feedback for enlarging speech movements in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 60*, 1818-1825. doi:10.1044/2017_JSLHR-S-16-0233

Yunusova, Y., Rosenthal, J. S., Rudy, K., Baljko, M., & Daskalogiannakis, J. (2012). Positional targets for lingual consonants defined using electromagnetic articulography. *The Journal of the Acoustical Society of America, 132*(2), 1027-1038. doi:10.1121/1.4733542

Yunusova, Y., Weismer, G., Kent, R. D., & Rusche, N. M. (2005). Breath-group intelligibility in dysarthria: Characteristics and underlying correlates. *Journal of Speech, Language and Hearing Research, 48*(6), 1294-1310. doi:10.1044/1092-4388(2005/090)

Yunusova, Y., Weismer, G., Westbury, J. R., & Lindstrom, M. J. (2008). Articulatory movements during vowels in speakers with dysarthria and healthy controls. *Journal of Speech, Language and Hearing Research, 51*(3), 596-611. doi:10.1044/1092-4388(2008/043)

APPENDIX

Training Stimuli

1. The girl wore her hair in two braids
2. The door slammed down on my hand
3. My shoes are blue with yellow stripes
4. The mailbox was bent and broken
5. I found a gold coin on the ground outside
6. The chocolate chip cookies smelled good
7. The church was white and brown
8. I went to the dentist the other day
9. The box was small and wrapped in paper
10. My pen broke and leaked blue ink
11. That guy has been talking forever
12. My daughter made the honour roll
13. I love a hot cup of coffee
14. Did you hear that song on the radio?
15. Don’t sit on the broken chair
16. The movie was coming out on videotape
17. He likes cheese and crackers for lunch
18. Could you please pass the jam
19. That’s my favourite Italian restaurant
20. Pick me up from the bank at eleven.
21. John planted the tree in the front yard.

22. He scored the winning touchdown
23. Sam loves the smell of fresh bread
24. Do you speak any other languages?
25. The family had their picture taken
26. The subway was running late tonight
27. We should have made a right turn.
28. How much does that chocolate cost?
29. The photographer is in the darkroom
30. He had a talent for writing music
31. She grows flowers in the greenhouse
32. Life in the country is relaxing
33. Jen adopted a new baby kitten
34. Ryan dropped his keys down the grate.
35. Remember to pay rent this month
36. Show me how to change the locks
37. The coat needs a new zipper
38. Luke went to college in England
39. We went on a road trip to Vegas
40. I always need my midnight snack
41. He'll clear the snow with a snow plow
42. Can we stop at the next gas station?
43. Have you seen my new painting?
44. Please don't stop telling me the story

45. Tell the neighbours to turn it down
46. Make a list before you go shopping
47. The plant needs more sun and water
48. Jimmy worked on a crossword puzzle
49. Using chopsticks is a real challenge
50. That was quite a strong argument.

Table 1.

Participant demographic and clinical characteristics.

ID	Age / Sex	PD Onset (y)	HY	Medication	MoCa (/30)	Previous Speech Therapy	Baseline SIT (%)	Baseline Scaled Speech Intelligibility (Z score)	Perceptual Characteristics
PD14	90 / M	5	2	Levodopa	25	None	94.55	-0.77	**Audible inspiration,** short phrases, voice stoppages, intermittent breathy voice, variable rate, monoloudness, monopitch, reduced stress
PD25	73 / M	2	1	Levodopa-carbidopa, pramipexole	30	None	96.36	-1.81	**Monopitch, monoloudness, imprecise consonants,** short rushes of speech, reduced stress, harsh voice, hypernasality
PD27	72 / M	3	1	Levodopa	29	None	96.36	-1.24	**Reduced stress, monopitch, monoloudness,** imprecise consonants, low pitch
PD28	77 / M	0.6	1	Levodopa	28	None	99.09	-1.13	**Increased rate overall, monoloudness, repeated phonemes**/ phrases**,** pitch breaks, breathy voice, short rushes of speech, monopitch, imprecise consonants
PD30	63 / M	4	-	Levodopa, pramipexole	30	4 years prior; details unknown	94.55	-2.14	**Imprecise consonants, repeated phonemes, breathy voice, monopitch, monoloudness,** short rushes of speech, pitch breaks, audible inspiration, increased rate overall, reduced stress

Note. PD = Parkinson's disease; HY = Hoehn and Yahr score; MoCA = Montreal Cognitive Assessment; SIT = Sentence Intelligibility Test; Scaled speech intelligibility scores are expressed as Z scores relative to healthy control speakers from larger study (Kearney et al., in press); Perceptual characteristics in bold were observed by both speech-language pathologists.

Table 2.

Baseline effect sizes showing participants' response to the verbal cue alone (uncued-cued).

Participant ID	**UNCUED-CUED effect size**
	Feedback
PD14	-0.54
PD25	1.05*
PD27	0.01
PD28	0.44
PD30	0.24

Note. * *d*>1.0, clinically significant difference from uncued to cued condition.

Table 3.

Mean and SD of percent change in AWS for cued sentences across treatment sessions with and without visual feedback.

Participant ID	**Percent Change AWS (SD)**	
	Feedback	**No Feedback**
PD14	83.84 (47.43)	37.10 (9.02)
PD25	280.95 (138.34)	55.79 (14.28)
PD27	172.83 (80.64)	56.63 (16.10)
PD28	46.46 (16.70)	27.34 (8.25)
PD30	769.55 (498.01)	82.62 (8.21)

Note. AWS = Articulatory working space.

Table 4.

Effect sizes pre-post treatment for cued-trained, cued-untrained, and uncued sentences.

Participant		**Pre-post effect size**	
ID	**Cued-trained**	**Cued-untrained**	**Uncued**
PD14	1.07*	0.28	0.04
PD25	1.13*	-0.53	-0.02
PD27	1.67*	1.26*	-0.33
PD28	0.40	-0.92	-0.45
PD30	1.64*	1.05*	-0.90

Note: * d>1.0, clinically significant change from pre-post treatment.

Table 5.

Percentage of paired-comparison ratings rated as being more intelligible or the same pre and post treatment for CUED-trained sentences. The values in bold indicate an improved intelligibility rating with larger movement size, recorded for PD28 following treatment.

Participant	**Percentage of Ratings**		
ID	**Pre**	**Post**	**Same**
PD14	70	10	20
PD25	75	0	25
PD27	30	30	40
PD28	15	**55**	30
PD30	50	25	25

Figures and Captions

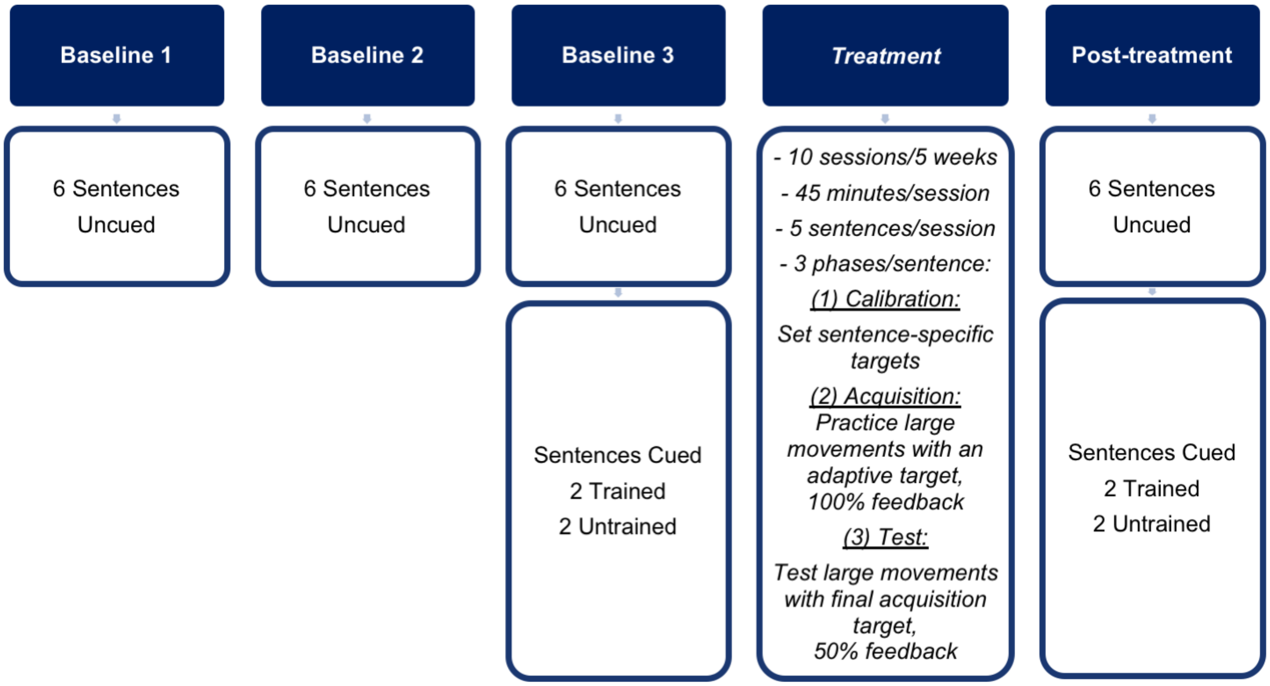

Figure 1. Flowchart of the assessment and treatment schedule.

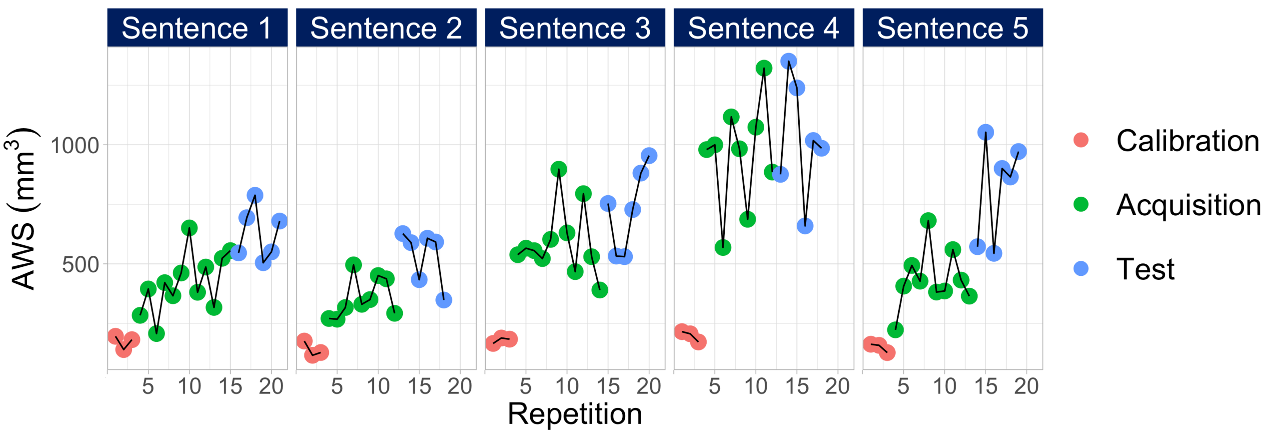

Figure 2. Sample data from a single treatment session for one participant (PD30).

Figure 3. Visual feedback in the form of two video games ("dragon world", "fish world") showed articulatory working space (AWS) of the tongue blade relative to a target AWS.

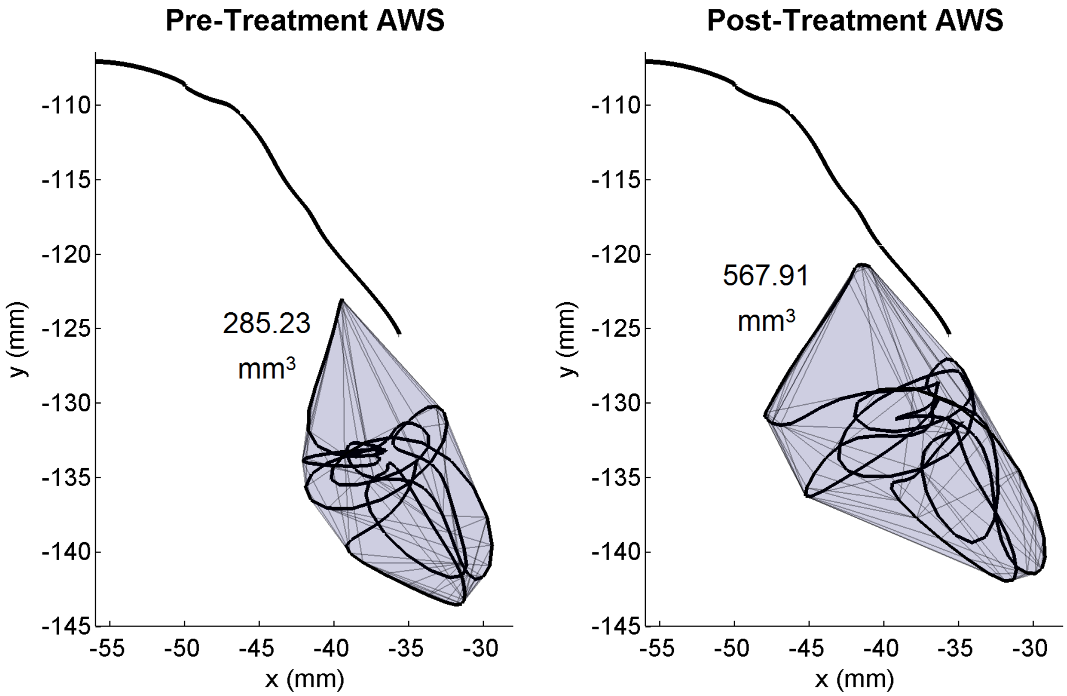

Figure 4. Articulatory working space (AWS) of the tongue blade trajectory obtained for speaker PD28 during the sentence "Jimmy worked on a crossword puzzle" displayed relative to the trace of the hard palate. In this example, a 99.1% increase in AWS was documented from pre to post treatment.

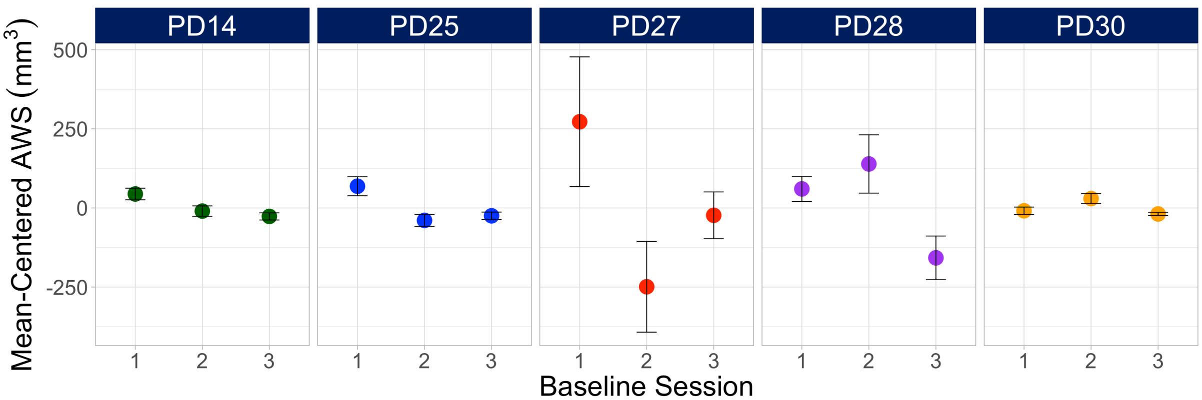

Figure 5. Mean and standard deviation of baseline measures of AWS for all uncued stimuli.

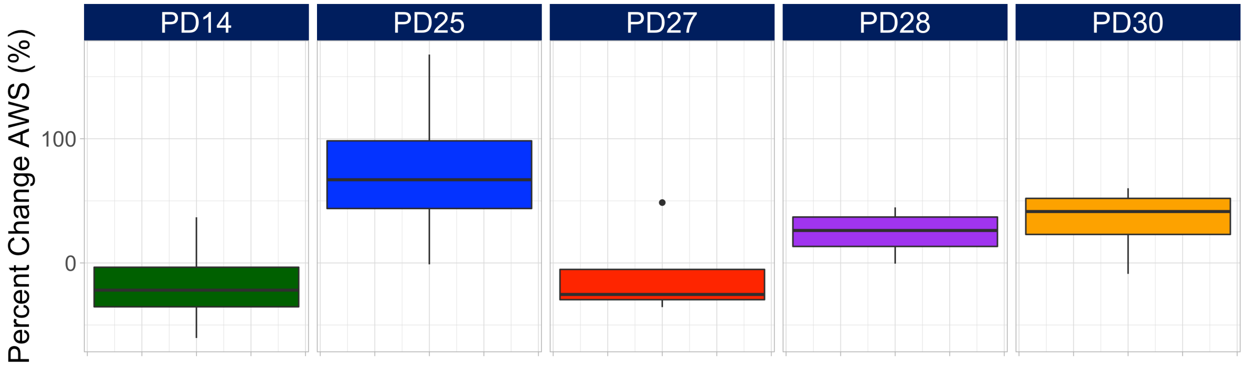

Figure 6. Box and whisker plots of percent change in AWS for cued (relative to uncued) stimuli.

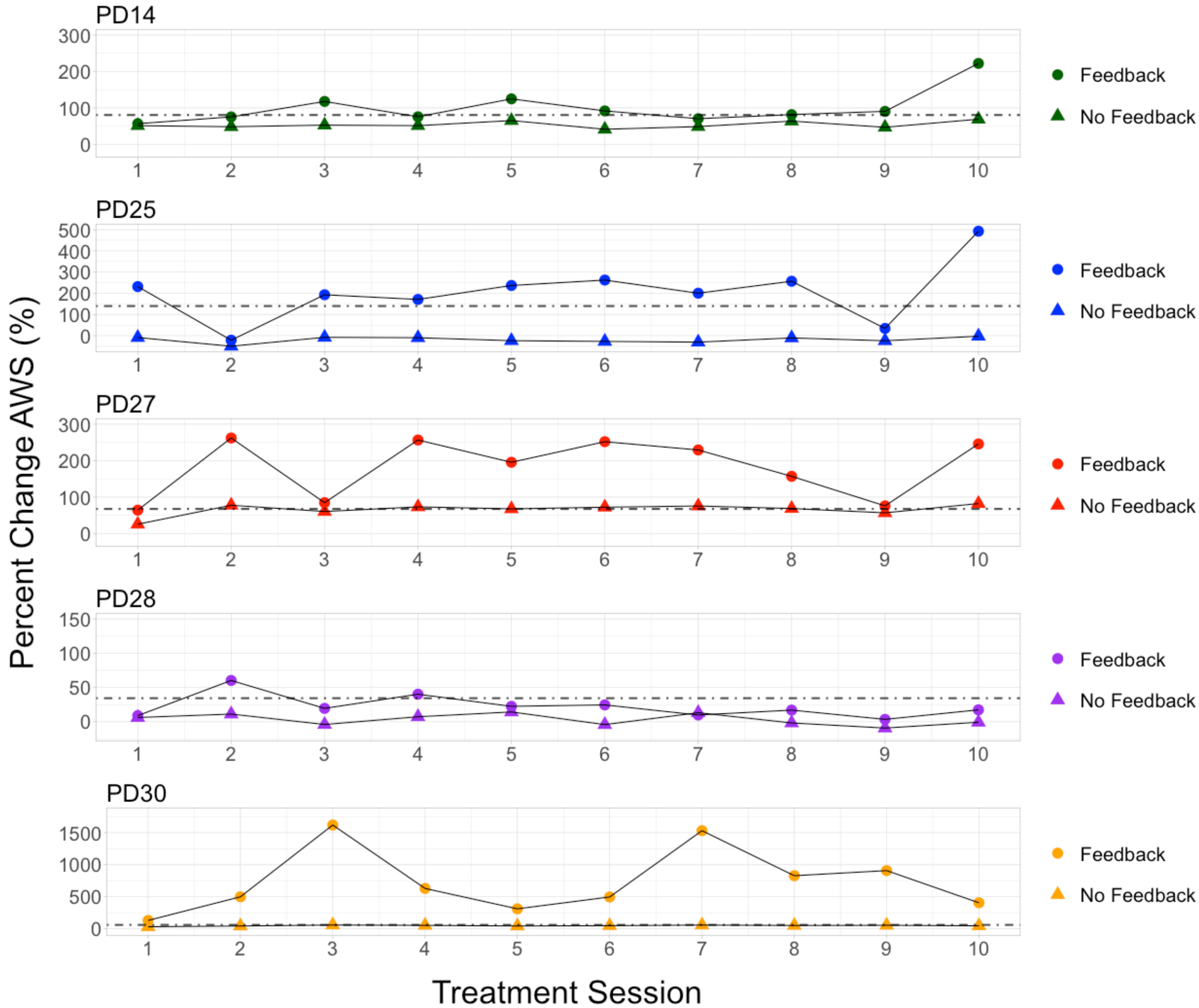

Figure 7. Percent change in AWS for cued sentences produced during treatment with and without visual feedback. The 'feedback' trials were trials following feedback and the 'no feedback' trials were those following no feedback. The dot-dashed line indicates 2SD above pre-treatment mean.