Innovative Technology for the Assisted Delivery of Intensive Voice Treatment (LSVT®LOUD) for Parkinson Disease

Angela E. Halpern, Lorraine O. Ramig, Carlos E. C. Matos, Jill A. Petska-Cable, Jennifer L. Spielman, Janice M. Pogoda, Phillip M. Gilley, Shimon Sapir, John K. Bennett, and David H. McFarland

Purpose: To assess the feasibility and effectiveness of a newly developed assistive technology system, Lee Silverman Voice Treatment Companion (LSVT® Companion™, hereafter referred to as “Companion”), to support the delivery of LSVT®LOUD, an efficacious speech intervention for individuals with Parkinson disease (PD).

Method: Sixteen individuals with PD were randomized to an immediate (n = 8) or a delayed (n = 8) treatment group. They participated in 9 LSVT LOUD sessions and 7 Companion sessions, independently administered at home. Acoustic, listener perception, and voice and speech rating data were obtained immediately before (pre), immediately after (post), and at 6 months post treatment (follow-up). System usability ratings were collected immediately post treatment. Changes in vocal sound pressure level were compared to data from a historical treatment group of individuals with PD treated with standard, in-person LSVT LOUD.

Results: All 16 participants were able to independently use the Companion. These individuals had therapeutic gains in sound pressure level, pre to post and pre to follow-up, similar to those of the historical treatment group.

Conclusions: This study supports the use of the Companion as an aid in treatment of hypokinetic dysarthria in individuals with PD. Advantages and disadvantages of the Companion, as well as limitations of the present study and directions for future studies, are discussed.

Key Words: Parkinson, speech treatment, technology

The great majority of individuals with Parkinson disease (PD) have voice and speech disorders (Hartelius & Svensson, 1994; Ho, Iansek, Marigliani, Bradshaw, & Gates, 1998; Logemann, Fisher, Boshes, & Blonsky, 1978). In spite of the deleterious effects of these disorders on communication and quality of life, only a small percentage of individuals in need receive speech therapy (Hartelius & Svensson, 1994).

Pharmacological, surgical, and traditional speech therapy methods have yielded disappointing results, in terms of magnitude of change and long-term effects of therapeutic outcome on voice and speech disorders associated with PD (Krause, Fogel, Mayer, Kloss, & Tronnier, 2004; Pinto et al., 2004; Rodriguez-Oroz et al., 2005; Trail et al., 2005). In contrast, research has shown that the Lee Silverman Voice Treatment (LSVT®LOUD) can produce significant and long-term improvements in voice and speech functions in individuals with idiopathic PD (see Pinto et al., 2004; Trail et al., 2005; Yorkston, Spencer, & Duffy, 2003). LSVT LOUD is an intensive, 1-month speech therapy program that trains individuals with PD to speak in a louder, good-quality voice and enables them to recalibrate their sensorimotor system through self-monitoring to increase vocal effort (Fox et al., 2006; Ramig et al., 1988). Although initially developed to improve voice and speech in individuals with PD, LSVT LOUD has been shown to be effective in the treatment of dysarthria in individuals with multiple sclerosis (Sapir et al., 2001), cerebellar dysfunction (Sapir et al., 2003), and stroke (Mahler, Ramig, & Fox, 2009) and, more recently, for pediatric populations such as those with cerebral palsy.
addition, cross-system improvements (McFarland & Tremblay, 2006) in nonspeech functions such as swallowing (El Sharkawi et al., 2002), tongue strength (Ward, Theodoros, Murdoch, & Silburn, 2000), and facial expression (Spielman, Borod, & Ramig, 2003) have been noted in preliminary data in individuals with PD following LSVT LOUD, even though these were not specific treatment targets. This implies that LSVT LOUD may target underlying neural processes and may have distributed effects on a variety of sensorimotor behaviors that are typically impaired in PD. The therapeutic protocol of LSVT LOUD corresponds to the growing body of basic and clinical neuroscience literature suggesting that intensive and functionally relevant exercise capitalizes on the neuroplasticity of impaired neural systems (Kleim & Jones, 2008; Kleim, Jones, & Schallert, 2003). This is reinforced by the results of preliminary brain imaging studies demonstrating changes in brain activation in response to LSVT LOUD (Liotti et al., 2003; Narayana et al., 2010).

The mode of delivery of LSVT LOUD addresses high-level neural deficits by incorporating many of the integral components necessary for motor learning, sensorimotor training, and neuroplasticity (Fox et al., 2006; Ramig, Fox, & Sapir, 2004). In particular, LSVT LOUD is delivered in a high dose that consists of four consecutive, individual 1-hr treatment sessions per week for 4 weeks (i.e., 16 sessions in 1 month), and homework practice is assigned for every day of the month in which a client is in treatment. Within treatment sessions, intensity is achieved through multiple repetitions of treatment tasks and through increased requirements for effort and accuracy as treatment sessions progress. This intense dosage has been identified as a key element in recent studies of neuroplasticity-principled approaches to exercise (Fox et al., 2006; Kleim & Jones 2008). To facilitate generalization of skills into functional communication, LSVT LOUD also targets the sensorimotor issues common in PD. This is accomplished by incorporating sensory calibration activities into the treatment and homework tasks. These activities train individuals to independently cue themselves to increase vocal amplitude to the proper loudness level and to feel comfortable that the “normal loudness” is not too loud. The maintenance of the therapeutic effects of LSVT LOUD even 2 years post-treatment (Ramig, Sapir, Countryman, et al., 2001) suggests that individuals who are treated intensively with LSVT LOUD learn to more appropriately scale vocal output independent of external cues, perhaps through reliance on internal cueing, auditory and kinesthetic feedback, and self-regulation.

Despite the established efficacy of LSVT LOUD (Ramig et al., 2004; Ramig, Sapir, Countryman, et al., 2001; Ramig, Sapir, Fox, & Countryman, 2001), the “real-world” treatment of speech and voice remains an unmet need for many individuals with PD. The present efficacious treatment dosage (16 individual 1-hr sessions in 1 month) may not be feasible for individuals who have mobility and/or geographical constraints, who are unable to obtain insurance reimbursement for speech treatment, or who lack access to trained and certified clinicians. In addition, there are not enough therapists to deliver treatment to all of the individuals who need it, and this issue will become even more exacerbated as the current population ages (Tanner & Goldman, 1996; ThinkQuest, n.d.; The Rand Corporation, 2001). These constraints are not unique to individuals with PD nor to LSVT LOUD. There is a need to find alternative solutions to increase treatment accessibility while maintaining treatment fidelity and preserving the elements essential for treatment success.

A potential solution is technology-supported intervention. Present technology has proven promising for various conditions, such as cognitive function and memory in chronic brain injury (Hart, Hawkey, & Whyte, 2002; Kim, Burke, Dowds, Boone, & Park, 2000; Kirsch et al., 2004), speech training for individuals with hearing loss (Massaro & Light, 2004), and ambulatory voice monitoring (Hillman, Heaton, Masaki, Zeitels, & Cheyne, 2006). Recently, telepractice has also been considered as a means to improve access to speech and voice evaluation and treatment (Hill et al., 2006; Mashima et al., 2003; Mashima & Doarn, 2008).

Applications of technology specifically to voice and speech in PD have also been reported (Constantinescu et al., 2010a). Several studies have demonstrated the feasibility of delivering LSVT LOUD via telepractice (Constantinescu et al., 2010b; Constantinescu et al., 2011; Howell, Tripoliti, & Pring, 2009; Theodoros et al., 2006; Tindall, Huebner, Stemple, & Kleiner, 2008). Speech enhancement, amplification (Cariski & Rosenbek, 1999; Greene, Watson, Gay, & Townsend, 1972), biofeedback (Rubow & Swift, 1985), and masking noise or delayed/altered auditory feedback (Adams & Lang, 1992; Adams et al., 2006; Downie, Low, & Lindsay, 1981; Hanson, & Metter, 1980) have also been explored as tools to improve voice or speech in PD and in Parkinson-related disorders. Other than telepractice delivery of LSVT LOUD (Constantinescu et al., 2011; Howell et al., 2009; Theodoros et al., 2006), applications of technology can induce immediate changes or improvement in voice and speech, but none of the technologies have documented or investigated long-term effects when the feedback or device is removed.

One potential solution to encourage and monitor long-term treatment gains is to use interactive computer programs that are emerging as powerful tools for clinical intervention for a variety of disorders and populations. They can be customized to specific treatment protocols and have the advantage of being used remotely by clients, thus overcoming many of the fundamental barriers to delivering intensive clinical protocols. It is for this reason that, in the present study, we investigated the application of the LSVT LOUD protocol delivered through an interactive, customized, personal digital assistant (PDA)–based software program. We chose to implement our software onto the PDA platform because of its portability and ease of use for individual home self-training. Our goals were (a) to explore a realistic and cost-effective means of increasing access to treatment and home practice and (b) to enhance and sustain treatment outcomes through technology-supported intervention.

All versions (initial and subsequent) of the software are named LSVT® Companion™ (hereafter, “Companion”). The Companion collects data on the variables trained in the LSVT LOUD treatment protocol (vocal sound pressure...
level [VocSPL], fundamental frequency [F0], and duration of phonation) and directs a client through an entire session of LSVT LOUD via interactive programming and audio and visual feedback. A detailed description and illustration of the program are provided in the Appendix.

The purposes of this study were to determine the feasibility of using the Companion as an independent “at-home” clinician for a portion of LSVT LOUD sessions, to determine if treatment gains achieved with the Companion-supported treatment are significant, and to determine if improvements in VocSPL are similar to gains achieved with standard in-person LSVT LOUD (Ramig, Sapir, Fox, & Countryman, 2001). The comparison to standard in-person LSVT LOUD is important in order to establish if Companion-supported therapy could be a viable addition to or alternative to in-person therapy. Data from a group of participants with PD who were treated with LSVT LOUD, as described in the Ramig, Sapir, Fox, and Countryman (2001) study, are included as the standard in-person comparison group (the historical treatment group; hereafter, “historical group”). If determined feasible, the findings from this study will have important implications for enhancing treatment accessibility not only for the individuals with PD but for all individuals who seek voice/speech treatment.

Method

Participants

Participants were recruited through local physicians, support groups, and aging centers as well as from advertisements in websites, newsletters, and facilities for aging individuals and individuals with PD. Twenty individuals with idiopathic PD initially entered the study; two individuals were excluded because of cognitive difficulties, one individual withdrew because of difficulties with travel, and one individual withdrew because of an unrelated medical problem. The data reported here are based on the 16 individuals who completed the study.

All individuals completed screening tasks, including the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh., 1975), the Beck Depression Inventory (BDI; Beck, Steer, & Brown, 1996), a voice/speech severity screening, an oral mechanism exam, a Companion device usability test (which assessed a participant’s ability to perform basic functions necessary to use the Companion), an audiometric screening, and a pre-treatment laryngeal exam performed by an otolaryngologist. All participants indicated their race as White and their ethnicity as not Hispanic or Latino.

The participants were diagnosed by their neurologist as having idiopathic PD and were considered “optimally medicated” at the start of the study. Two participants had a change in PD medication from both pre to post treatment and post to 6-month follow-up, and a third participant had a change from post to 6-month follow-up. Participant comments (available in three out of five instances) indicated that the change in medication did not affect voice and speech. Upon examining the data, the change in medication had no obvious effects.

Participants did not have severe depression or dementia (as determined by the BDI and the MMSE); did not show symptoms of another neurological condition other than or in addition to PD; had no drug abuse (indicated by a neurologist and/or medical records); had no neurosurgeries, head injuries, or head or neck cancer; had no significant history of gastrointestinal disease or surgery (defined by an otolaryngologist); had no history of smoking within the last 4 years; had no speech or voice disorders that were unrelated to PD; or lack of evidence of a voice or speech disorder (defined by a speech-language pathologist [SLP]). One participant scored a 28 on the BDI (severe = 29–63). Although inclusion was defined as a score of 0 to 24, because this participant indicated problems in the areas of sleep patterns and appetite (areas commonly affected by PD and PD medications), it was determined that he would be an appropriate candidate. Six of the 16 participants had tremor in their dominant hand, and three had apparent upper body dyskinesias. The participants were not involved in any other form of speech treatment during the time of this study, had not participated in speech treatment within the last 4 years, and had not previously participated in LSVT LOUD. All participants had hearing that was within normal limits for their chronological age. Information regarding computer experience was available for 12 participants, all of whom indicated that they had at least minimal computer experience.

Participants were randomized to an immediate treatment group (hereafter, “immediate group”; n = 8) or a delayed treatment group (hereafter, “delayed group”; n = 8). Refer to the Procedures and Data Collection for Outcome Measures section for a description of the study schedule for each group. To achieve group assignment balance on age, gender, years since diagnosis, stage of PD (Hoehn & Yahr, 1967), and voice and speech severity, we randomized using a minimization approach with a random component—that is, once the minimization algorithm determined the optimal treatment assignment for a given participant, the probability that the participant was actually assigned to that treatment was .90. For the single variable of VocSPL, the immediate group and the delayed group were compared to a group of individuals with idiopathic PD (n = 13) from a previous study (Ramig, Sapir, Fox, & Countryman, 2001) who were randomly assigned to a treatment group. This historical treatment group (hereafter, “historical group”) received standard in-person LSVT LOUD (all 16 sessions delivered by an expert LSVT LOUD clinician, who was the third author of that study). Data collection and treatment for the historical group occurred in a similar timeframe and manner as data collection and treatment in the present study.

Treatment

All participants in the immediate group and the delayed group received intensive voice treatment (LSVT LOUD) modified through use of the Companion. All treatment sessions in the present study were provided and directed by three clinicians specializing in LSVT LOUD (the first, fourth, and fifth author). The schedule and mode of treatment administration were based on previous treatment studies that have determined optimal treatment dosage (Ramig, Countryman, Thompson, & Hori, 1995): four consecutive days a week for 16 individual 60-min sessions in 1 month.
The only variation to this schedule was that the participants were present for up to an additional 30 min for downloading and analyzing data from the Companion and providing additional training or explanation regarding the use of the device as needed. Carryover and homework assignments followed the same intensity and frequency as those given during standard in-person LSVT LOUD. All clinicians followed the LSVT LOUD protocol for in-person sessions. Companion training and instructions for at-home treatment were developed by the same clinicians and were conducted with participants according to a pre-established protocol to provide consistency with instruction. Participants were given instruction sheets and were asked to fill out home session treatment logs to facilitate and track compliance.

LSVT LOUD tasks were implemented using the Companion as follows. Participants completed in-person LSVT LOUD in the clinic for nine of the 16 sessions and used the Companion for therapy independently at home for seven treatment sessions. In Week 1, participants worked with the clinician on all 4 days and were trained on the Companion during Day 4 of treatment. Participants were also asked to watch a video over the weekend that reviewed the training information and provided further instruction. In Week 2, participants came to the clinic for treatment on Day 1. The Companion was used during this session. The participants were asked to use the device as independently as possible. Care was taken to ensure that all participants were practicing voice exercises using healthy voice quality (Ramig et al., 1995). The participants then used the Companion at home on Day 2, worked with the clinician on Day 3, and used the Companion at home on Day 4. In Week 3, the participants worked with the clinician on Day 1 and then used the Companion at home for the remaining three sessions of the week. In Week 4, participants worked with the clinician on Day 1, with the Companion on Days 2 and 3, and with the clinician on Day 4. The Companion was occasionally used during the in-person session to determine levels for goal setting and to check the accuracy of data collected. VocSPL, F0, and duration data were collected by the Companion as the participants practiced in the clinic and on their own. These data were downloaded by the speech clinician at each meeting, progress was evaluated, and new goals were programmed into the Companion as appropriate. The at-home Companion session data as well as the homework and carryover assignments were reviewed at each in-person session. In this way, if a participant did not correctly complete the activities or completed them on the wrong day, review and additional instruction could happen in a timely manner. In addition, participants were asked for qualitative feedback about the usefulness and challenges of the device each time they met with the clinician.

**Procedures and Data Collection for Outcome Measures**

We analyzed four outcome measures to assess the impact of the Companion: (a) acoustic analysis of pre, post, and follow-up outcome data; (b) listener perception ratings; (c) ratings of voice and speech by participants with PD and their significant other; and (d) ratings of usefulness of the Companion by participants and their significant others. See Table 1 for a summary of the outcome measures.

**Outcome measure 1: Acoustic.** Pre, post, and follow-up treatment data were collected on six different days for the immediate group (twice pre-treatment [pre 1 and pre 2], twice post-treatment [post 1 and post 2], and twice 6 months following treatment [follow-up 1 (FU 1) and follow-up 2 (FU 2)]) and on eight different days for the delayed group (four times pre-treatment [pre 1, pre 2, pre 3, and pre 4], twice post-treatment [post 1 and post 2], and twice 6 months following treatment [FU 1 and FU 2]). The delayed group participated in pre 1 and pre 2, followed by 1 month of no treatment, and then participated in pre 3 and pre 4, followed by 1 month of treatment. Thus, the delayed group served as a control group during their no-treatment month. The immediate group participated in pre 1 and pre 2, followed by 1 month of treatment. Pre 1 and pre 2 (immediate group) and pre 3 and pre 4 (delayed group) were collected the week prior to the treatment phase of the study. All post data were collected the week following the treatment phase of the study. All attempts were made to schedule each participant’s data collection sessions at the same time of the day across all data collection sessions. Treating clinicians were not visible to participants during post-treatment data collection and did not elicit data from a

**TABLE 1. Summary of outcome measures.**

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Effect following treatment?</th>
<th>Variables</th>
<th>Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective acoustic data</td>
<td>Increases in VocSPL</td>
<td>VocSPL obtained at pre, post, and 6 months following treatment</td>
<td>Reading, sustained “ah,” monologue, picture description, verbal fluency</td>
</tr>
<tr>
<td>Naïve listener perception</td>
<td>Perceived differences between pre- and post-treatment speech</td>
<td>Numeric ratings ranging from –50 (worse) to +50 (better) for pre/post sentence pairs</td>
<td>Sentence pairs presented at both normalized and restored VocSPL</td>
</tr>
<tr>
<td>Self and other perception</td>
<td>Perceived changes in voice and speech</td>
<td>Ratings of voice and speech using standard questionnaires</td>
<td>Self: VHI Other: CETI-M, VAS</td>
</tr>
<tr>
<td>Companion usability</td>
<td>Usability and helpfulness of the Companion</td>
<td>Ratings on a 1–5 scale of Companion “helpfulness”; response to questionnaire</td>
<td>Companion questionnaires completed post treatment by self and other</td>
</tr>
</tbody>
</table>

**Note.** VocSPL = vocal sound pressure level; VHI = Voice Handicap Index (range: 0–120; higher ratings reflect more difficulty); CETI-M = Communicative Effectiveness Index—Modified (range: 0–10; higher ratings reflect better performance); VAS = Visual Analogue Scale (range: 0%–100%; higher ratings reflect better performance).
participant whom they treated. Data for the historical group from the Ramig, Sapir, Fox, and Countryman (2001) study were collected on seven different days (three times pre treatment [pre 1, pre 2, and pre 3], twice post treatment [post 1 and post 2], and twice approximately 6 months following treatment [FU 1 and FU 2]). The timing of the recording periods relative to treatment in the historical group was similar to the timing in the immediate group in the present study. With the exception of the pre 1 and pre 2 recordings that occurred prior to the 1 month of no treatment for the delayed group, the timing of the pre 3, pre 4, post, and FU recordings were also similar for the delayed group.

Acoustic data for the immediate group and the delayed group were collected in an Industrial Acoustics Corporation (IAC) sound-treated booth via a head-mounted AKG-410 microphone placed 8 cm from each participant’s lips. In order to extract VocSPL directly from the microphone signal, we calibrated the microphone using a Type I sound-level meter (Brue&Kjaer [B&K] Model 2238) at a distance of 30 cm using standard procedures (Svec, Popolo, & Titze, 2003). We recalibrated the microphone in the event that it was readjusted for distance or if any amplifiers were changed during the course of the recording.

The speech tasks analyzed in the present study were as follows:

1. Sustaining vowel phonation for maximum duration (hereafter, Max Ah task).
2. Reading “The Rainbow Passage” (Fairbanks, 1960; hereafter, Rainbow task).
3. Speaking spontaneously while describing a picture (hereafter, Picture task).
4. Speaking spontaneously while describing the “happiest day they can recall” (hereafter, Monologue task).

In the historical study, Fluency was not included, and Monologue differed slightly in that participants were asked to talk about a topic of their choice. Each task was performed once on each recording day, except for Max Ah, which was performed six times.

**Acoustic measurements.** VocSPL was measured for all tasks. We initially edited sound files in the present study to remove any noise that was not speech. We then extracted average VocSPL from the calibrated microphone signal using a customized software program developed by the third author (Matos, 2003). In this program, VocSPL was extracted through use of an automatic algorithm that takes four readings of acoustic RMS power per second, selects the peak of these four values, and returns the average of the selected peak values together with the SD for the whole sound sample. These peaks are comparable to VocSPL measurements from a standard Type I sound-level meter. To determine the reliability of this method, we randomly selected 30 original sound files, each containing a specific task, to be reanalyzed by a different data analyzer. Data collection and measurement procedures for the historical group are described in Ramig, Sapir, Fox, and Countryman (2001).

**Outcome measure 2: Listener perceptual ratings.** Four listeners (ages 24, 27, 30, and 33 years) rated pre/post sentence pairs for the listener perception task. Pre 2 (PRE) and post 1 (POST) samples were used for the immediate group. Pre 4 (PRE) and post 1 (POST) samples were used for the delayed group. These recording days were considered to afford the most conservative comparison between pre and post sessions. One listener was an SLP; one had just received a master’s degree in speech-language pathology, and the other two were speech-language pathology graduate students. These listeners were made aware that the speakers had participated in various types of speech treatment; however, the listeners were naive to the speakers’ diagnosis and treatment assignment. All listeners reported hearing within normal limits and having English as their first language.

The following sentences, which were taken from “The Rainbow Passage” (Fairbanks, 1960), were included in the study:

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow.

The rainbow is a division of white light into many beautiful colors.

There is, according to legend, a boiling pot of gold at one end.

Sentence material was selected to be long enough to provide a connected speech sample and yet short enough to provide a uniform, consistent sample within and across speakers.

Listeners completed both a training task and the study via computer, which they could control. The computer screen displayed a line that contained a “–50” on the left end and a “+50” on the right end. For each pair of sentences, the listeners were instructed to listen to Sample A followed by Sample B. Listeners were then asked to decide if Sample B sounded better or worse than Sample A and to rate how much better (1 to 50) or worse (–1 to –50), with –50 being much worse, +50 being much better, and 0 being the same. “Better” and “worse” were defined as global measures including voice quality, articulatory clarity, rate, and intonation.

Listeners completed the study, one listener at a time, in a sound-treated booth. Two Bose Companion®2 speakers were positioned 60 cm from one another and 22 cm from the edge of the table, on either side of the computer screen. After the initial training, listeners were presented with 60 paired samples (16 participants × three sentence pairs + 12 pairs selected at random to test intrarater reliability) played at a restored VocSPL and with 60 paired samples played at a normalized VocSPL. Thus, listeners rated a total of 120 paired samples. The randomly selected clips for reliability analyses included a few instances of two pairs of clips from the same participant. In these cases, one clip pair was eliminated so that only one clip pair from each participant was included in the final analysis. Thus, 10 paired clips were used for intrarater agreement and reliability, and 10 clips were used for interrater agreement and reliability for both normalized and restored conditions.

Listeners were reminded to double-check their answers to ensure that they included the “negative” sign for ratings.
that they intended to be –50 to –1. However, in reviewing the data, one intrarater reliability clip pair for Listener 4 and two clip pairs for Listener 2 were obvious outliers (i.e., they were scored in the opposite direction by 30 or more points from 0) from the other clip pairs for that same participant. Thus, the validity of these clip pairs was questionable, and they were not included in the final analysis.

We used restored VocSPL to assess the direct treatment effect of targeting increased VocSPL to improve voice and speech. We used normalized VocSPL to investigate treatment-related changes in voice and articulation separate from perceived loudness. We used custom-built MATLAB programs to normalize and restore speech samples. We completed normalization by amplifying or attenuating a group of sound files to match a common arbitrary level of 70 dB. Restored samples were converted back to their original VocSPL at 30 cm. We checked normalized and restored samples using a custom-built SPL extractor (Matos, 2003). Any sample that deviated from its original or the standard by more than 1 dB was hand corrected using Adobe Audition Version 1.5.

Paired samples were randomized for PRE–POST order of presentation. Scores from clips rated in the POST–PRE order were converted to PRE–POST scores by multiplying by –1. In addition, two listeners heard the 60 normalized samples first, and two listeners heard the 60 restored samples first. Speech samples were presented at a “comfortable” listening level as determined by two SLPs. This level was set prior to the first listener session and was maintained for all listeners.

Outcome measure 3: Ratings of voice and speech by participating and their significant others. To evaluate the impact of treatment-related changes on the participants and their significant others, we administered standard voice and speech rating scales. Participants completed self-ratings using the Voice Handicap Index (VHI; Jacobson et al., 1997). The VHI contains 30 statements (e.g., “My voice is worse in the evening”) that assess the functional, physical, and emotional impact of an individual’s voice difficulties.

Participants’ close friends or family members (e.g., significant others) completed a Visual Analogue Scale (VAS; Ramig et al., 1995; Spielman, Halpern & Ramig, 2010) and an early, 10-question version of The Modified Communication Effectiveness Index (CETI–M; Ball, Beukelman, & Pattee, 2004). The VAS requires significant others to rate the participants’ voice and speech based on the percent in time in which they feel that particular voice and speech characteristics are present (0%–100%). The CETI–M asks significant others to indicate how effective they believe that the participant is at communicating in each of 10 different situations.

Outcome measure 4: Ratings of usefulness of the Companion by participants and their significant others. User involvement and feedback are key elements in human-centered design (International Organization for Standardization [ISO], 1999) for computer-based systems (Cole, 2006). Thus, in addition to evaluating speech outcome data, it is essential to determine if participants can independently work with the Companion and if they find it helpful. To obtain information on the usefulness of the Companion, usability methods were employed with participants and families via question-asking protocols (i.e., rating forms and interviews; Usability First, n.d.).

### Statistical Analyses for Participant Baseline Characteristics and Outcome Measures

**Participant baseline characteristics.** We conducted data analysis and comparisons among the immediate group, delayed group, and historical group using raw data for all three groups. We used an analysis of variance (ANOVA) to test for differences by treatment group in the baseline characteristics of age, years since diagnosis, voice and speech severity, and PD stage, excluding those with unavailable data. We used Fisher’s exact test for the characteristic of gender. Tests were two sided, with significance levels of .05.

**Outcome measures 1, 2, and 3.** In general, we used paired $t$ tests to examine the differences between two measurements of the same variable recorded during a single treatment phase (i.e., pre, post, or follow-up) and for intrarater agreement.

For listener perception ratings, we computed quantitative agreement as the absolute difference in scores, with level of agreement categorized as 0–5, 6–10, 11–20, and > 20. Choice of cutpoints was guided by the distributions of quantitative agreement as a continuous variable. Intrarater agreement for each listener was assessed by the percentage of PRE/POST sentence pairs, for which agreement was within 5, 10, and 20 points. Overall intrarater agreement was calculated as the mean percentage of sentence pairs at each of these agreement levels. We assessed interrater agreement by calculating mean percentage of pairwise agreement within 5, 10, and 20 points. Comparisons on mean agreement levels between restored and normalized sentence pairs were done using $z$ tests. We used mixed-model ANOVAs to estimate intraclass correlation coefficients (ICCs) for intra- and interrater reliability; specifically, ICC was defined as the estimate of subject variance divided by the estimate of total variance (Shrout & Fleiss, 1979). We derived 95% confidence intervals (CIs) using bootstrapped estimates of the standard errors of the ICCs (Efron & Tibshirani, 1986). For each of 1,000 bootstrap samples, ratings from a given listener for 10 participants randomly sampled with replacement were used to calculate the intrarater ICC for that listener and sample. Overall intrarater reliability was calculated as the mean ICC across all four listeners. For each of 1,000 bootstrap samples, ratings from all four listeners for 10 participants randomly sampled with replacement were used to calculate the interrater ICC for that sample. ICC standard errors were estimated to be the standard deviations of the 1,000-sample ICCs. ICCs between restored and normalized sentence pairs were compared through the use of $z$ tests.

We used mixed-model ANOVAs to test for differences in VocSPL, VHI, VAS, and CETI–M by fixed-effects treatment and period, with study participant nested within treatment as a random effect. Initial models included a term for the interaction between treatment and period. Final models excluded the interaction term if it was nonsignificant; otherwise, stratified analyses were performed. Pairwise comparisons were done through the use of linear contrasts. We used mixed-model ANOVAs to test for differences in listener perception ratings by fixed-effect sentence pair type (restored vs. normalized); study participant and sentence pair number (1–3), nested within participant and sentence pair type, were random effects.
All tests were two sided, with significance levels of .05. No adjustments were made for initial multiple comparisons between recording periods for acoustic data (e.g., pre 1 vs. pre 2). Šidák–Bonferroni and Tukey–Kramer corrections were applied to multiple comparisons between pre, post, and FU sessions for acoustic data and for self-ratings and other ratings, respectively. We performed analyses using SAS® Version 9.0 (SAS Institute, Cary, NC) and NCSS 2007 (NCSS LLC, Kaysville, UT).

**Results**

**Participant Baseline Characteristics**

The immediate group, delayed group, and historical group were similar at baseline in age, gender, time since diagnosis, and speech and voice severity. Stage of PD (Hoehn & Yahr, 1967) was similar for the immediate group and delayed group. Stage data were not available for the historical group. Individual baseline characteristics can be found in the Supplemental Table. Group mean participant data are summarized in Table 2.

**Outcome Measure 1: Acoustics**

Among all participants combined, VocSPL pre 1 and pre 2 were similar. There were no significant differences between post 1 and post 2 or FU 1 and FU 2 in either the immediate group or the delayed group. In the delayed group, pre 3 and pre 4 were also similar, but the mean of pre 1 and pre 2 was significantly lower than the mean of pre 3 and pre 4 for Fluency task and Picture task, respectively (by 0.98 dB and 0.73 dB, respectively). On the basis of these results, by-subject pre-treatment scores (PRE) were defined as the mean of pre 1 and pre 2 for the immediate group (one participant had only pre 2 available; thus, PRE was based on one measurement for this participant) and the mean of pre 3 and pre 4 for the delayed group. In the historical group, by-subject PRE was defined as the mean of pre 1, pre 2, and pre 3 (or the mean of pre 1 and pre 2 for one subject with only the first two pre recordings available; and the mean of pre 2 and pre 3 for one subject for the Monologue task). For all three treatment groups, POST and FU were defined as the mean of post 1 and post 2 and the mean of FU 1 and FU 2 (only one FU was available for one of the subjects in the historical group), respectively.

In analysis of interrater agreement, VocSPL did not significantly differ between original and repeat sound files. VocSPL is summarized in Table 3. For all tasks, the Treatment (immediate group, delayed group, historical group) × Period (PRE, POST, FU) interaction was nonsignificant. In analyses of main effects, period but not treatment was highly significant ($p < .0001$ for each task). The strongest effects were observed between PRE and POST and between PRE and FU ($p < .0001$ for all pairwise comparisons). This effect was a result of an increase in VocSPL from the PRE to the POST condition, as indicated by average difference scores, which ranged from 5.7 dB to 17.3 dB across tasks and groups. There was a weak effect from POST to FU, as indicated by a decrease in VocSPL, which ranged from −0.4 dB to −3.7 dB across tasks and groups. However, a strong significant effect was still maintained PRE to FU resulting from an increase in VocSPL, as indicated by average difference scores that ranged from 2.0 dB to 15.2 dB across tasks and groups.

**Outcome Measure 2: Listener Perceptual Ratings**

Both intra- and interrater agreement within 20 points was substantial (around 90% for both types of agreement) and did not differ by whether the clips were restored or normalized. Intrarater agreement within 5 points was somewhat better in normalized than in restored clips ($p = .06$). The mean intrarater ICC was .69 (95% CI [.60, .77]) for restored clips and .78 (95% CI [.64, .93]) for normalized clips ($p = .02$). The mean interrater ICC was .66 (95% CI [.46, .86]) for restored clips and .60 (95% CI [.36, .84]) for normalized clips ($p = .70$).

The overall mean converted post-treatment listener rating score was 19.8 ($SD = 14.3$), on the −50 to +50 scale. Significantly greater improvement was perceived by listeners in the restored condition (actual VocSPL), $M = 22.5$ ($SD = 13.4$), compared with the normalized condition, $M = 17.1$ ($SD = 14.6$). All scores were significantly different from 0 ($p < .0001$).

**Outcome Measure 3: Ratings of Voice and Speech by Participants and Their Significant Other**

Ratings between sessions in the same recording period for VHI, VAS, and CETI–M revealed no significant differences among available data between pre 1 and pre 2 for the

---

**TABLE 2. Mean (SD) and range of descriptive data for participants in each of three groups.**

<table>
<thead>
<tr>
<th>Group</th>
<th>Age in years</th>
<th>Years since diagnosis</th>
<th>Hoehn &amp; Yahr stage$^a$</th>
<th>Voice and speech severity$^b$</th>
<th>Gender (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>Range</td>
<td>M (SD)</td>
<td>Range</td>
<td></td>
</tr>
<tr>
<td>Immediate (n = 8)</td>
<td>65.8 (8.4)</td>
<td>56–81</td>
<td>4.4 (3.5)</td>
<td>1–11</td>
<td>1.9 (0.4)</td>
</tr>
<tr>
<td>Delayed (n = 8)</td>
<td>63.3 (6.0)</td>
<td>54–74</td>
<td>4.7 (3.1)</td>
<td>1–11</td>
<td>2.0 (0.7)</td>
</tr>
<tr>
<td>Historical (n = 13)</td>
<td>68.5 (9.0)</td>
<td>51–80</td>
<td>8.5 (6.5)</td>
<td>1.5–20</td>
<td>2.7 (1.6)</td>
</tr>
</tbody>
</table>

Note. Immediate = immediate treatment group; Delayed = delayed treatment group; Historical = historical treatment group. Em dashes indicate data not available.

$^a$Hoehn & Yahr stages are 1–5, with higher stages indicating greater severity. $^b$Severity ratings of speech and voice deficits are on a scale of 0–5 (0 = none, 1 = mild, 5 = severe).
### TABLE 3. Mean (SD) vocal sound pressure level (VocSPL) from PRE, POST, and FU for individual tasks and groups.

<table>
<thead>
<tr>
<th>Task</th>
<th>PRE&lt;sup&gt;a&lt;/sup&gt; M (SD)</th>
<th>POST&lt;sup&gt;a&lt;/sup&gt; M (SD)</th>
<th>FU&lt;sup&gt;b&lt;/sup&gt; M (SD)</th>
<th>PRE to POST significance p</th>
<th>PRE to FU significance p</th>
<th>POST to FU significance p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rainbow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate group</td>
<td>67.4 (1.1)</td>
<td>76.5 (3.6)</td>
<td>74.8 (4.1)</td>
<td>&lt; .0001</td>
<td>&lt; .0001</td>
<td>.0006</td>
</tr>
<tr>
<td>Delayed group</td>
<td>68.8 (2.0)</td>
<td>76.8 (3.2)</td>
<td>74.2 (2.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical group</td>
<td>71.5 (3.2)</td>
<td>78.4 (3.8)</td>
<td>75.4 (3.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max Ah</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate group</td>
<td>67.7 (3.6)</td>
<td>85.0 (2.4)</td>
<td>82.9 (4.5)</td>
<td>&lt; .0001</td>
<td>&lt; .0001</td>
<td>ns</td>
</tr>
<tr>
<td>Delayed group</td>
<td>70.4 (5.4)</td>
<td>85.2 (6.3)</td>
<td>83.4 (2.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical group</td>
<td>68.9 (6.3)</td>
<td>82.9 (3.7)</td>
<td>79.2 (4.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate group</td>
<td>67.2 (1.4)</td>
<td>73.0 (3.6)</td>
<td>72.6 (3.3)</td>
<td>&lt; .0001</td>
<td>&lt; .0001</td>
<td>.005</td>
</tr>
<tr>
<td>Delayed group</td>
<td>66.3 (2.0)</td>
<td>73.4 (3.1)</td>
<td>71.0 (2.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical group</td>
<td>69.3 (3.6)</td>
<td>75.0 (3.8)</td>
<td>71.3 (4.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate group</td>
<td>66.4 (0.9)</td>
<td>74.0 (4.0)</td>
<td>72.7 (3.5)</td>
<td>&lt; .0001</td>
<td>&lt; .0001</td>
<td>ns</td>
</tr>
<tr>
<td>Delayed group</td>
<td>65.9 (2.4)</td>
<td>73.8 (3.0)</td>
<td>71.5 (3.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical group</td>
<td>69.0 (4.8)</td>
<td>74.9 (4.1)</td>
<td>72.6 (4.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate group</td>
<td>68.2 (1.2)</td>
<td>76.5 (3.8)</td>
<td>75.2 (3.7)</td>
<td>&lt; .0001</td>
<td>&lt; .0001</td>
<td>ns</td>
</tr>
<tr>
<td>Delayed group</td>
<td>68.1 (1.9)</td>
<td>75.9 (3.8)</td>
<td>73.7 (3.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical group</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. PRE = immediately before treatment; POST = immediately after treatment; FU = follow-up (6 months after treatment). Fluency was not collected for historical group. p values reflect period effect by task.

<sup>a</sup>Immediate, n = 8; Delayed, n = 8; Historical, n = 13. <sup>b</sup>Immediate, n = 8; Delayed, n = 8; Historical, n = 7.

---

immediate group (except for VAS “always loud”; p = .03), pre 3 and pre 4 for the delayed group, or post 1 and post 2 sessions. FU 1 and FU 2 differed for VHI (p = .004) but not for VAS or CETI–M. The overall correlations of time 1 versus time 2 were high (.96), even in the two instances of a significant difference. This suggests that the averages of time 1 and time 2 are representative overall and are not correlated to time. Thus, in order to provide the best representation of the data, and to more accurately account for the participant variability that is inherent in perceptual ratings, when available, we averaged the data from both sessions within the same recording period. Data from repeat sessions were available an average of 98.5% (SD = 3.0) of the time for the VHI, 57.7% (SD = 8.2) of the time for the five tasks from the VAS, 56.5% (SD = 11.5) of the time for the CETI–M total score, and 58% (SD = 8.6) of the time for the three specific CETI–M items.

**Self-rating: VHI.** The Treatment (immediate group, delayed group × Period (PRE, POST, FU) interaction was nonsignificant for the VHI. A change of 18 points represents a significant shift in psychosocial function at the individual level, per Jacobson et al. (1997). Of the 16 participants, three demonstrated a positive shift of this magnitude from PRE to POST. Although the score declined at FU for these three participants, one participant maintained a positive shift at FU, and three additional participants demonstrated a positive shift from PRE to FU.

**Significant other: VAS.** The following items were analyzed: “always loud enough,” “never a hoarse voice,” “never monotone,” “always speaks so others can understand,” and “always starts a conversation.” The Treatment × Period interaction was significant for PRE versus POST (immediate group, delayed group, POST versus FU (immediate group), and PRE versus FU (delayed group). The Period × Item interaction was also significant PRE versus POST for four items and was significant PRE versus FU and POST versus FU for “always loud” (p < .05).

**Significant other: CETI–M.** Analysis of CETI–M included CETI–M total score (CETI–M total) and three selected CETI–M ratings (CETI–M select) of effectiveness that are commonly reported as difficult (Dykstra, 2010): “travelling in a car” (“car”), “speaking in a noisy environment” (“noisy”), and “speaking in a long conversation” (“long conversation”). The Treatment × Period interaction was significant for PRE versus POST (immediate group and delayed group, CETI–M total; immediate group, CETI–M select), POST versus FU (immediate group, CETI–M select), and PRE versus FU (delayed group) (Tukey–Kramer [95% CI] ps < .05). The Period × Item interaction was also significant PRE versus POST for all measures (ps < .01) and PRE versus FU for CETI–M total, “noisy,” and “long conversation” (ps < .05).

Specific results for PRE, POST, and FU comparisons of period by items for VHI, CETI–M, and VAS are summarized in Table 4.

**Outcome Measure 4: Ratings of Usefulness of the Companion by Participants and Their Significant Others**

All 16 participants reported that they were able to use the Companion independently at home. The participants responded to questions regarding their experience with the
TABLE 4. PRE, POST, and FU mean (SD), and significance for participant (self) and significant other (SO) ratings of voice and speech.

<table>
<thead>
<tr>
<th>Task</th>
<th>Period M (SD)</th>
<th>Period pairwise contrast p values*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRE n</td>
<td>POST n</td>
</tr>
<tr>
<td>VHI: Self</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate group</td>
<td>8</td>
<td>45.6 (20.9)</td>
</tr>
<tr>
<td>Delayed group</td>
<td>8</td>
<td>43.1 (14.9)</td>
</tr>
<tr>
<td>VAS: SO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never monotone: Immediate</td>
<td>8</td>
<td>62.6 (21.7)</td>
</tr>
<tr>
<td>Never monotone: Delayed</td>
<td>8</td>
<td>64.1 (22.9)</td>
</tr>
<tr>
<td>Never hoarse: Immediate</td>
<td>8</td>
<td>67.6 (22.0)</td>
</tr>
<tr>
<td>Never hoarse: Delayed</td>
<td>8</td>
<td>64.9 (18.4)</td>
</tr>
<tr>
<td>Always loud: Immediate</td>
<td>8</td>
<td>45.9 (16.6)</td>
</tr>
<tr>
<td>Always loud: Delayed</td>
<td>8</td>
<td>50.3 (22.8)</td>
</tr>
<tr>
<td>Always starts a conversation: Immediate</td>
<td>8</td>
<td>67.3 (16.6)</td>
</tr>
<tr>
<td>Always starts a conversation: Delayed</td>
<td>8</td>
<td>71.4 (19.3)</td>
</tr>
<tr>
<td>Always speaks so others can understand: Immediate</td>
<td>8</td>
<td>58.0 (18.7)</td>
</tr>
<tr>
<td>Always speaks so others can understand: Delayed</td>
<td>8</td>
<td>61.8 (22.9)</td>
</tr>
<tr>
<td>CETI–M: SO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SO total rating: Immediate</td>
<td>8</td>
<td>71.1 (14.2)</td>
</tr>
<tr>
<td>SO total rating: Delayed</td>
<td>8</td>
<td>63.9 (19.8)</td>
</tr>
<tr>
<td>Traveling in car: Immediate</td>
<td>8</td>
<td>6.6 (2.0)</td>
</tr>
<tr>
<td>Traveling in car: Delayed</td>
<td>8</td>
<td>6.1 (2.3)</td>
</tr>
<tr>
<td>Noisy environment: Immediate</td>
<td>8</td>
<td>5.2 (1.4)</td>
</tr>
<tr>
<td>Noisy environment: Delayed</td>
<td>8</td>
<td>4.8 (2.8)</td>
</tr>
<tr>
<td>Long conversation: Immediate</td>
<td>8</td>
<td>6.4 (2.4)</td>
</tr>
<tr>
<td>Long conversation: Delayed</td>
<td>8</td>
<td>5.3 (1.9)</td>
</tr>
</tbody>
</table>

Note. For the Voice Handicap Index (VHI) self ratings, lower scores indicate less severity. For the Visual Analogue Scale (VAS) significant other (SO) ratings, scores are based on percent of time that voice or speech behavior was observed (0%–100%). For the Modified Communication Effectiveness Index (CETI–M) SO ratings, 1 = not at all effective, 10 = extremely effective.

*p < .05 using Tukey–Kramer (95% confidence interval) correction for multiple comparisons.

Discussion

We designed this study to determine the feasibility of using a newly developed, interactive assistive technology system, the Companion, as an independent “at-home” clinician for a portion of LSVT LOUD therapy. We also designed this study to determine if treatment gains achieved with Companion-supported therapy were significant and similar to gains achieved by a historical treatment group (“historical group”) who received standard in-person LSVT LOUD.

The results of the study demonstrate that the participants made significant gains PRE to POST and PRE to FU treatment in VocSPL. These gains in VocSPL were comparable to those obtained by the historical group with standard in-person LSVT LOUD (Ramig, Sapir, Fox, & Countryman, 2001). The results also indicate that the participants were able to independently use the Companion at home, and the vast majority rated the device as very helpful. In addition, improvements in voice and speech were identified by listeners, as demonstrated by positive overall mean POST treatment rating scores in both the restored and normalized VocSPL conditions. This suggests that the listeners perceived not only changes in VocSPL but also changes in other factors of speech and voice that contribute to a perception of “better,” such as voice quality and articulation, as indicated by positive ratings in the normalized condition as well. Finally, perceptual assessment by significant others further indicated favorable outcomes with the Companion.

Although many treated participants reported informally that the treatment had a positive impact on their communication and they described powerful improvements in quality of life, this was not reflected in the group VHI results. The reasons for this discrepancy are not clear. Ho, Bradshaw, and Iansek (2000) compared individuals with PD and hypophonic dysarthria with healthy age- and gender-matched control individuals with normal voice and speech. They found that although the individuals with PD spoke more quietly than did the controls, they nevertheless perceived (immediate and playback perception) their own speech to be louder than the controls. Therefore, it is possible that in the present study, the individuals with PD were much less aware of the magnitude of their hypophonia prior to receiving LSVT LOUD, and as a result of education during treatment, they became increasingly more aware of their voice characteristics. Thus, in spite of improvements in VocSPL from PRE to POST, this
increased awareness could contribute to the lack of a significant difference between the PRE and POST VHI ratings.

In the present study, clinicians administered the LSVT LOUD during nine sessions of in-person training; the remaining seven sessions were then completed by the participants independently using the Companion. One potential concern of any “independent” vocal exercise without direct clinician oversight is that clients may generate vocal targets with hyperfunctional behaviors. The results of the listener perception study and preliminary observations of group pre/post videostroboscopy data (in preparation) do not support this concern. These findings and impressions are consistent with previous reports indicating that LSVT LOUD does not induce vocal strain or hyperfunction (Countryman, Hicks, & Ramig, 1997; Smith, Ramig, Dromey, Perez, & Samandari, 1995) and instead results in improved voice quality (Baumgartner, Sapir, & Ramig, 2001). These results suggest that many participants were able to use the techniques taught during sessions with the clinician to achieve a good-quality louder voice when practicing alone.

As mentioned previously, an important component of LSVT LOUD is training individuals to internalize the sense of effort required to achieve a voice that is perceived by others to be at a normal loudness level. In other words, the treatment encourages sensorimotor recalibration of appropriate vocal effort and voice quality. One potential worry is that participants would rely on the external cues provided by the Companion and that treatment effects would not generalize beyond this specific cueing. To mitigate this reliance on cueing, questions were built into the Companion that required the participants to think about and rate the amount of effort they were using during the exercises. In addition, participants were allowed to use the Companion only for their 1-hr at-home LSVT LOUD session; they were not allowed to use the Companion for their homework or carryover activities. This was done to encourage participants to learn to internally generate the right amount of effort for a louder voice that is perceived by others as within normal limits. Our findings—that the participants improved VocSPL and sustained this improvement at 6 months post-treatment (without the use of the Companion during the post-treatment period)—suggest that they indeed learned to independently produce a voice of more normal loudness than they had produced prior to therapy.

Another potential concern is that our older individuals with PD would have difficulty using technology due to lack of experience, fear of technology, and arm/hand tremor or upper-body dyskinesia. However, post-treatment questionnaires indicated that the participants enthusiastically welcomed and used the Companion. Even participants up to the age of 81, with little computer experience and with arm/hand tremor and upper body dyskinesia, were able to complete the study.

The use of the Companion in clinical intervention offers the opportunity for enhanced access to LSVT LOUD by removing many of the common barriers to treatment—such as mobility and geographical constraints, lack of LSVT-certified clinicians, and the cost of standard in-person treatment—thus maintaining more individuals in need of treatment while maintaining the desired frequency and intensity of practice.

In addition, with continued use following treatment, the Companion may be especially powerful for long-term maintenance and carryover. Consistent with the principles of motor learning, skill acquisition, muscle training, and neuroplasticity, all of the Companion tasks are practiced multiple times with consistent encouragement to improve performance with each trial (Fox et al., 2006; Kleim & Jones, 2008; Ramig et al., 1995). The Companion may facilitate increased frequency of practice and enhance trial-by-trial performance by providing online knowledge of results (feedback) to motivate the individual and reinforce targeted vocal behaviors. This knowledge of results is especially important for individuals with PD, as they tend to rely—much more so than individuals with normal neurological function—on external feedback to master a newly learned motor skill (Guadagnoli, Leis, Van Gemmert, & Stelmach, 2002).

Another advantage to using the Companion is the ability to precisely measure the targeted vocal behaviors. This is important for training, clinical accountability, and research and teaching purposes. Moreover, the technology developed for the LSVT LOUD program can be applied to other treatment programs and different voice and speech disorders.

**Limitations to the Study**

There are some potential limitations of the present study, including the small number of individuals with PD who were studied and the small number of listeners who participated in the perceptual studies. Because of the exploratory nature of the study, numerous comparisons were presented, and p values should be interpreted accordingly.

In this study, we screened individuals with PD and limited participation to those without severe cognitive impairment, without severe signs of clinical depression, with relatively mild or moderate PD, and with adequate hearing. This sample is unlikely to be representative of the entire population of individuals with PD and, thus, generalization of these results may be limited.

Some participants described difficulty achieving an optimal voice during home practice and required more frequent instruction regarding proper voice production. A limitation to the present study is the inability to determine exactly what type of voice these participants were producing during home practice sessions.

**Future Directions and Clinical Implications**

We are presently investigating improvements to our technology platform that could improve the fidelity of treatment received with the Companion and could address some of the present study limitations. These improvements include a recording mechanism that would allow individuals to send “at-home” samples of their voice to their supervising SLP and would allow individuals to listen to their own voice for real-time calibration feedback. Other enhancements include (a) modifying voice analysis subroutines for the Companion to analyze and provide feedback regarding “acceptable” versus “unacceptable” qualitative aspects of the voice and (b) expansions to other technology platforms.
Future work is needed to explore different treatment regimens with different client characteristics in order to more thoroughly understand applicability to a broader treatment context. In addition, it will be important to investigate changes in facial expression, articulation, and swallowing following use of the Companion to determine if these results are similar to those reported following standard in-person LSVT LOUD (Dromey, Ramig, & Johnson, 1995; El Sharkawi et al., 2002; Sapir, Spielman, Ramig, Story, & Fox, 2007; Spielman, et al., 2003). These findings would help to guide recommendations for treatment delivery. It may be determined that individuals with more severe impairments require standard in-person LSVT LOUD, in contrast to the use of assistive computer technology. In fact, we envision that the Companion will be one tool to be used along a continuum of care from in-person to technology-supported “independent” treatment, on the basis of disease severity and other client characteristics.

In conclusion, the present results suggest that assistive technology for the delivery of LSVT LOUD may show great promise for overcoming many of the constraints to treatment delivery while maintaining treatment efficacy and impact.

Acknowledgments

This research was funded by National Institute on Deafness and Other Communication Disorders and National Institute of Neurological Disorders and Stroke Grant R21-DC05583, The Michael J. Fox Foundation, and The Coleman Institute for Cognitive Disabilities. We thank our research assistants Heather Gustafson and Leslie Mahler. We also express our gratitude to the individuals who volunteered their time to participate in this study.

References


Appendix
Companion Device Description and Screen Shot of One Exercise

The Compaq iPAQ3650 and 3800 personal digital assistants (PDAs) were selected for this study on the basis of their multimedia capabilities, processing power, and development tools (see Figure A-1). The heart of the application is a module that monitors incoming audio signals and computes both vocal sound pressure level (VocSPL) and fundamental frequency (F0) in real time. This information is relayed to the other modules, where it receives specific processing. We calibrated the VocSPL readings with a Bruel & Kjaer Type 1 sound level meter (SLM), and we detected F0 via an autocorrelation-based algorithm (Matos, 2001, 2002). Each PDA was individually calibrated for VocSPL, resulting in an average VocSPL difference between the SLM and PDA of 0.7 dB ($\text{SD} = 0.5$) for manual and autocalibration. When compared to a vocal demodulator (Winholtz WVD-100), F0 was detected with an average error of 2.0 Hz (range = 0–4), a negligible error for pitch perception and for monitoring clinically significant changes in F0. Because of the beta nature of the program, we made some changes to the software during the course of the study to improve device calibration, accuracy of SPL measurement, and general functionality. Only changes that were operational in nature and would not differentially impact participant performance were allowed.

The LSVT®LOUD Companion™ was designed to guide an individual through an entire LSVT®LOUD session while collecting data on VocSPL and F0. These data are recorded to a file that the clinician can download and analyze. Targets for VocSPL, F0, and duration were individualized for each participant and were revised by the treating clinician as often as needed. Goals could be easily set or updated by clinicians simply typing in the target VocSPL, F0, and duration into boxes on a dedicated goal-setting screen on the Companion. We developed the interface so that it is simple to use for both clinicians and participants while carefully considering the specific visual, motor, cognitive, and voice/speech needs of individuals with PD. Participants started the program by touching the word “LSVT,” which appeared in the Start menu. Once the program began, all instructions for how to control and participate in the exercises were provided via the Companion. We designed the backgrounds in bright, contrasting colors that have common meaning (e.g., green means “go”) to make the targeted goals easier to identify. In addition, the visual displays use concrete, familiar objects (e.g., thermometer, piano, clock) that are designed to be intuitive and require little interpretation. Because the device is intended to be used without a clinician, feedback was designed to emulate what would be received from a clinician in a therapy session. Thus, in addition to the written comments shown on the screen, a wide variety of auditory feedback phrases such as “Good job!”, “Let’s get louder,” “Great!”, “Can you go any higher?”, and short instructions at the beginning of each exercise were incorporated. A major component of LSVT LOUD is training individuals to monitor their vocal effort. Therefore, exercises in the Companion incorporate a task that requires participants to rate their level of effort following each exercise.

Participants were allowed to complete their 1-hr at-home session at a time that was convenient for them. However, they were instructed that sessions should occur on particular days and should be completed in a 1-hr stretch. Review of home session data allowed clinicians to determine if participants had correctly completed the session or if further instruction was required.

FIGURE A-1. Screen shot of one exercise.