

Summer 2015

# Speech function in persons with Parkinson's disease: effects of environment, task and treatment

Carrie E. Rountrey  
*James Madison University*

Follow this and additional works at: <https://commons.lib.jmu.edu/diss201019>



Part of the [Speech Pathology and Audiology Commons](#)

---

## Recommended Citation

Rountrey, Carrie E., "Speech function in persons with Parkinson's disease: effects of environment, task and treatment" (2015).  
*Dissertations*. 32.  
<https://commons.lib.jmu.edu/diss201019/32>

This Dissertation is brought to you for free and open access by the The Graduate School at JMU Scholarly Commons. It has been accepted for inclusion in Dissertations by an authorized administrator of JMU Scholarly Commons. For more information, please contact [dc\\_admin@jmu.edu](mailto:dc_admin@jmu.edu).

Speech Function in Persons with Parkinson's Disease: Effects of Environment, Task  
and Treatment

Carrie Ellen Rountrey

A dissertation submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

Communication Sciences and Disorders

August 2015

## **Dedication**

This work is dedicated to two amazing, academic women who I feel are still with me. To my grandmother, Alma Simmons Rountrey, always encouraged and supported my academic pursuits. When she said that no answer to any question was ever good enough for me, and that I couldn't sit still or stay in one place for long, the teacher in her meant it as a compliment, even when the grandmother in her missed me. And to my friend and former colleague, Dr. Marion Meyerson, who opened my eyes and my mind to new ways of looking at things, taught me that skepticism and an open mind coexist, and thought of me as a scientist from the time we met. I strive to live, learn and love the way these women did each day, personally and professionally.

“One never notices what has been done, one can only see what remains to be done.” – Marie Curie

“We do not grow absolutely, chronologically. We grow sometimes in one dimension, and not in another; unevenly. We grow partially. We are relative. We are mature in one realm, childish in another. The past, present, and future mingle and pull us backward, forward, or fix us in the present. We are made up of layers, cells, constellations.” - Anaïs Nin

## **Acknowledgements**

Many practitioners talk about bridging theory to practice, and after 12 years of practice, this project has allowed me to bridge back to the theory, and to question practice through my own research. My objective for my time at JMU was to become a scientist and a skeptical investigator of clinical practice. This journey is far from over, and I am grateful to those who prompted the journey to begin.

My advisor, Dr. Christy Ludlow has been enthusiastic about my research questions and helped to refine my ideas into research with clinical application and benefit to the professional community. Dr. Rory DePaolis was ever the supporter as I moved through stages of my research and advised on methodology and technology. Dr. Cristina Kuo continuously pointed me back to the literature, reminding me to look at where my current research questions fit, and inspired questions for the future through our discussions. Dr. Cynthia O'Donoghue, worked to see that my process was as seamless as possible. Dr. Jessica Huber was always available to dialogue about the project and ideas stemming from it. Each of your contributions brought this work to life.

The community of those affected by Parkinson's disease supported this work through participation, spreading the word, and allowing me to be a small part of their lives for a short time. I am inspired by their thirst for knowledge and dedication to research.

My fellow doctoral students walked alongside me at this time. Sharing this road with you made it less treacherous and more enjoyable! I have the special privilege of counting two of my cousins amongst my friends and colleagues. Thanks to Dr. Jonathan Greenberg, for support and help with algorithms to analyze intelligibility; and Dr. Amy Throckmorton, who gave me feedback on my very first prospectus draft, when I was afraid to send it to anyone else. My friends and former colleagues, Dr. Wendy Quach, Prof. Jean Jackson, Dr. Gloria Weddington, Dr. Henriette Langdon, and Dr. Aliaa Khidr have stayed in touch and encouraged me along my way. Last but of course, not least, my gratitude to my family, Mom Cathy and Dad Bill, brother, Josh, my son Owen, and my fiancé, J.R. Gentle. You've seen the good, the bad, and the ugly, and have loved me throughout. Thank you for sacrificing your time to give me time for this.

## Table of Contents

<b>Dedication .....</b>	<b>ii</b>
<b>Acknowledgements .....</b>	<b>iii</b>
<b>List of Tables .....</b>	<b>v</b>
<b>List of Figures .....</b>	<b>vi</b>
<b>Abstract .....</b>	<b>vii</b>
<b>Introduction.....</b>	<b>1</b>
<b>Research Questions .....</b>	<b>7</b>
<b>Significance .....</b>	<b>9</b>
<b>Innovation .....</b>	<b>11</b>
<b>Approach.....</b>	<b>13</b>
<b>To determine the effect of the environment on speech intensity and intelligibility     in persons with PD and healthy controls. ....</b>	<b>13</b>
Methods. ....	14
<b>To explore which clinical task predicts speech intelligibility in the natural     environment. ....</b>	<b>31</b>
Methods. ....	34
<b>To assess the effect of treatment on speech intelligibility and intensity in the     natural environment.....</b>	<b>40</b>
Methods. ....	42
<b>Potential Limitations .....</b>	<b>44</b>
<b>Manuscript: Effects of Environment and Task on Speech Functioning in Parkinson’s Disease and Healthy Controls.....</b>	<b>63</b>
<b>Pilot Study: Application of Ambulatory Recording Device for Measurement of a 12-week Treatment Protocol.....</b>	<b>103</b>
<b>Conclusions .....</b>	<b>125</b>

## **List of Tables**

<b>Table 1: Data, format, measures and analysis for environment study .....</b>	<b>20</b>
<b>Table 2: Coefficients of variation for pilot data for environment study.....</b>	<b>30</b>
<b>Table 3: Data, format, measures and analysis for exploratory study .....</b>	<b>35</b>
<b>Table 4: Data, format, measures and analysis for treatment study.....</b>	<b>43</b>
<b>Table 5: Pilot data for treatment study .....</b>	<b>44</b>
<b>Tables from Manuscript.....</b>	<b>94</b>
<b>Table 1: Participant Clinical Data .....</b>	<b>96</b>
<b>Table 2: Examination of Means.....</b>	<b>97</b>
<b>Tables from Pilot.....</b>	<b>122</b>
<b>Table 1: Participant Clinical Data .....</b>	<b>122</b>

## List of Figures

<i>Fig. 1 Power analysis using pilot test data on 4 PD patients on difference in speech intelligibility .....</i>	<b>14</b>
<i>Fig. 2 Using pilot test data on 4 PD patients on change in sound pressure level</i>	<b>15</b>
<i>Fig. 3 Basic linear interpolation graph for calibrating the recording devices ...</i>	<b>19</b>
<i>Fig. 5 Analyzing intensity of a sentence via Praat .....</i>	<b>25</b>
<i>Fig 5a. Select the sentence to be analyzed from the beginning to the end of the sound contour .....</i>	<b>26</b>
<i>Fig 5b. Select “intensity” and “get intensity,” or read intensity in green from the right side of the screen.....</i>	<b>27</b>
<i>Fig. 6 Pilot data for environment study .....</i>	<b>29</b>
<i>Fig. 7 Measuring <math>SPL_{VOWEL}</math> .....</i>	<b>38</b>
<i>Fig. 8 Measuring maximum phonation time .....</i>	<b>39</b>
<i>Fig. 9 Goberman &amp; Coelho 2002 disease factors and response fluctuations .....</i>	<b>45</b>
<i>Figures from Manuscript .....</i>	<b>98</b>
<i>Figure 1: LENA-DLP .....</i>	<b>98</b>
<i>Figure 2: 5b. Data Processing for Intensity .....</i>	<b>99</b>
<i>Figure 3: Examination of Intensity Between Groups.....</i>	<b>100</b>
<i>Figure 4: Spontaneous Speech Percent Intelligibility by Environment and Group....</i>	<b>101</b>
<i>Figure 5: Regression of Sentence Intelligibility Test on Percent Sentence Intelligibility in the Home .....</i>	<b>102</b>
<i>Figures from Pilot .....</i>	<b>123</b>
<i>Figure 1: The SpeechVive wearable device for treatment of hypophonia in Parkinson’s Disease .....</i>	<b>123</b>
<i>Figure 2: The LENA-DLP.....</i>	<b>124</b>

## **Abstract**

Parkinson's Disease (PD) is a degenerative neurological disease affecting aspects of movement, including speech. Persons with PD are reported to have better speech functioning in the clinical setting than in the home setting, but this has not been quantified. New methodologies in ambulatory measures of speech are emerging that allow investigation of non-clinical settings.

The following questions are addressed: Is speech different between environments in PD and in healthy controls? Can clinical tasks predict speech behaviors in the home? Is treatment proven effective by measures in the home? What can we glean from methods of measurement of speech function in the home?

The experiment included 13 persons with PD and 12 healthy controls, studied in the clinical and home environments, and 7 of those 13 persons with PD participated in a treatment study.

Major findings included: Spontaneous speech intelligibility, not intensity, was the differentiating factor between persons with PD and healthy controls. Intelligibility and intensity were not related. Both groups presented with higher sentence intensity in the home environment. Spontaneous speech intelligibility in the clinic was related to spontaneous speech intelligibility in the home. The Sentence Intelligibility Test emerged as the best predictor of spontaneous speech intelligibility in the home. Differences between pilot treatment groups measured in the home on intensity and intelligibility were not large enough to make a clinical trial feasible. Individual differences may account for many of these results, for example more severely impaired patients may have shown different data.



Drawing conclusions regarding the home environment via measures outside the home should be carefully considered. Ambulatory measures of speech are a viable option for studying speech function in non-clinical settings, and technology is advancing. Further investigation is needed to develop methodologies and normative values for speech in the home.

## **Introduction**

Parkinson's disease (PD) is a neurological disorder that involves the loss of dopamine producing cells in the basal ganglia, with disease progression to affect the cortex (Braak et al., 2003; Del Tredici, Rub, de Vos, Bohl, & Braak, 2002). In 2005, between 4.1 and 4.6 million people worldwide over the age of 50 were estimated to have PD. This number was projected to more than double to between 8.7 and 9.3 million by the year 2030 (Dorsey et al., 2007). Symptoms of PD include resting tremor, reduced range of movement (rigidity) and slow movement (bradykinesia) all of which affect functions such as walking, speech and writing (Gelb, Oliver, & Gilman, 1999). As these symptoms progress, the affected person may have difficulty walking, talking, or completing simple tasks. Studies of limb movement show that persons with PD have motor execution difficulties due to rigidity and bradykinesia (DeLong & Georgopoulos, 2011; Morris, Iansek, Matyas, & Summers, 1994; Stelmach, Teasdale, Phillips, & Worringham, 1989). All of the speech subsystems (respiratory, laryngeal, and supralaryngeal) can be affected in persons with PD (Connor, Abbs, Cole, & Gracco, 1989; Duffy, 2012; Forrest, Weismer, & Turner, 1989; Hammer & Barlow, 2010; Huber & Darling, 2011; Sadagopan & Huber, 2007). Environmental cues have been shown to improve motor movements in PD (Ho, Bradshaw, Iansek, & Alfredson, 1999; Oliveira, Gurd, Nixion, Marshall, & Passingham, 1997)

Eighty-nine percent of those with PD have voice problems and 45% have articulation problems (Logemann, Fisher, Boshes, & Blonsky, 1978). Persons with PD patients suffer from hypokinetic dysarthria, which includes reduced voice loudness, monoloudness, monopitch, disordered rate, disordered articulation,

hoarse, breathy, or harsh voice and/or tremulous voice (Darley, Aronson, & Brown, 1969b; Logemann et al., 1978).

Several objective and subjective measures of hypophonia and dysarthria are included in a standard speech and voice evaluation for PD and related disorders (Hillman, Montgomery, & Zeitels, 1997; Roy et al., 2013). The preferred practice patterns for Speech–Language Pathologists (American Speech-Language-Hearing Association, 2004) for voice and motor speech assessments include the measurement of: mean sound pressure level (SPL) in a prolonged vowel, range of SPL (SPL- $\Delta$ ), the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) (Kempster, Gerratt, Abbott, Barkmeier-Kraemer, & Hillman, 2009; Zraick et al., 2011), Sentence Intelligibility Test (SIT) (Yorkston, Beukelman, & Tice, 1996), maximum phonation time (MPT), harmonic to noise ratio (HNR), fundamental frequency ( $F_0$ ), and range of  $F_0$  ( $F_0$ - $\Delta$ ). Clinicians will also often include quality of life scales such as the Voice Handicap Index (VHI) (Jacobson et al., 1997). Speech disorders that accompany PD impair intelligibility (Duffy, 2012). Estimates of speech intelligibility are often used to determine if a neurological disease affects the speech mechanism and to index the severity of the impairment in everyday life (R. Kent et al., 1989).

Treatment approaches for dysarthria in PD (i.e. rate reduction, increased voice loudness, and “clear speech” usually involving increased effort and/or overarticulation) extend across an entire utterance and impact multiple speech components (i.e. respiration, phonation, articulation, resonance) (Yorkston, Hakel, Beukelman, & Fager, 2007). These treatment techniques are intended to improve

speech intelligibility (Duffy, 2012; Fox, Morrison, Ramig, & Sapir, 2002; Weismer, Yunusova, & Bunton, 2012; Yorkston et al., 2007) by improving multiple speech components, including phonation. Therapy techniques that instruct patients to use “loud” and “clear” speech show improvement in intelligibility in the clinical setting (Ramig, 1992; K. Tjaden, Sussman, & Wilding, 2014). The Lee Silverman Voice Therapy *LOUD* program (LSVT-*LOUD*) uses the sole instruction of “think loud,” to focus on voice loudness, which impacts multiple speech components simultaneously and increase speech intelligibility (Fox et al., 2002; Ramig, 1992).

Treatment approaches that include environmental cues, such as the wearable SpeechVive device, are also being developed. Instead of telling the patient explicitly to increase their speech loudness, the SpeechVive delivers cocktail chatter to the patient through one ear and employs the Lombard effect so that the patient will speak more loudly in the presence of the background noise (Pick & al., 1989; Stathopoulos et al., 2014). Principles of practice and exercise physiology indicate that heavier loading on the muscles results in increased strength (Taaffe, Duret, Wheeler, & Marcus, 1999). Currently Huber and colleagues are evaluating whether 12 weeks use of the Speech Vive increases speech intensity and intelligibility (Stathopoulos et al., 2014). The study measured patients’ intensity and intelligibility in their natural environment, pre- and post- 12 weeks of SpeechVive™ treatment. Patients’ speech intelligibility and intensity were measured when not wearing the device.

Improved intelligibility is the primary aim of treatment; however, this has not been measured as a result of speech therapies for PD (Deane et al., 2002).

Intelligibility is defined as “the degree to which the speaker’s intended message is recovered by the listener,” (R. Kent et al., 1989). When judging intelligibility within the clinical context (speech therapy office, rehab setting, laboratory), Hawthorne, behavioral and task effects should be considered (Franke & Kaul, 1978; McCarney et al., 2007a; Wickström & Bendix, 2000). Reading tasks such as the SIT, allow patients to concentrate on *how* they are speaking, while in natural communication settings patients must formulate *what* they are going to say in addition to *how* they are speaking (Connor & Abbs, 1991; Kempler & Lancker, 2002; Yorkston, 2010).

Even in controlled environments, the effect of task can be seen. In a single case study involving a male with PD, 5 tasks in the clinic were transcribed for a percentage of intelligibility (spontaneous speech, repetition, reading, repeated singing, spontaneous singing), and intelligibility in spontaneous speech (29%) was significantly lower than in the other four tasks (78-88%) (Kempler & Lancker, 2002).

The primary interest of this work is a comparison of two distinctly defined environments. For the purposes of this project, consider “in the clinic,” to mean any controlled clinical or lab environment. These may include outpatient treatment rooms, inpatient hospital beds, university laboratories, etc. Structured speech tasks are elicited in the clinic environment, and spontaneous speech may take place as well. The “natural environment,” refers to those places in which the participants would normally find themselves. These may include home, church, a friend or family member’s house, the market, the workplace, etc. Structured speech tasks are not elicited in the natural environment; spontaneous conversational speech is the

item of interest here. More specifically, the “home environment” includes that environment where the participants were communicating with their primary communication partner(s) in their most familiar environment, and in reference to this project, “clinical environment,” is the described laboratory where testing took place.

Differences in speech produced in the clinic versus in the natural environment in spontaneous situations have been reported anecdotally and in the literature for speakers with PD (Keintz, Bunton, & Hoit, 2007; Sarno, 1968; Weismer, 1984). Patients tend to focus on a single task, speech production, when performing nearly any evaluative task in the clinical environment. However, in the natural environment attention is divided from speech production in order to coordinate the demands of speech production with those of functional communication and other activities, causing many stimuli to compete for resources (Dromey & Benson, 2003). Keitz et al. (2007) observed in their pilot speakers with PD a substantial performance effect (judged by authors and spouses of participants) where patients would be much more intelligible during clinical tasks than was typical during natural environment conversation. This is consistent with findings by Sarno (1968) and Weismer (1984). Sarno delineated between “clinical performance” and “functional performance,” much in the same ways as this project delineates between the clinical and natural environments. She found it “unfortunate” that the standard of practice was to exclusively measure speech impairments using clinical performance, as there was such a disparity between clinical and functional speech performance, “especially true in patients who have Parkinson’s disease.” Years later,

Weismer (1984) noted the same in a study he conducted with patients with PD, that when reviewing the intelligibility of those with PD in the clinical setting versus “spontaneous” situations, the majority of the subjects were much more intelligible when producing experimental sentences than they were when engaged in spontaneous speech. Keitz et al (2007) attempted to synthesize more naturalistic speech in their full study by introducing a dual task, in this case, subjects were asked to read sentences while screwing in a bolt with their non-dominant hand. Ho et al. (2002) also used this strategy, as well as Dromey and Benson (2003) with healthy participants.

In this project, we seek to study actual spontaneous speech within the natural environment, appreciating these researchers’ efforts to synthesize naturalistic speech with dual tasks. Motivation, environment and speech materials are influences on speech production and resulting intelligibility (Hustad & Weismer, 2007), which encompass the majority of the clinical evaluation. Discreetly measuring intelligibility and intensity within the natural environment may reduce Hawthorne, behavioral and task effects on outcomes. Questions remain as to the relationship between measurements in highly controlled conditions and more realistic conditions (R. D. Kent, Weismer, Kent, & Rosenbek, 1989). To assess intelligibility in the natural environment and quantify potential therapeutic effects, measures should be made in the natural environment or in a clinical simulation of a natural communication environment. This project seeks to overcome these clinical influences and remaining questions by studying the natural environment itself.

## Research Questions

A series of research questions are outlined on measurement of speech function of persons with PD and healthy controls within the natural environment, with the natural environment being their functional everyday environment (home, work, church, etc.). Hypophonia accompanying PD is characterized by a soft voice and reduced intelligibility (Darley, Aronson, & Brown, 1969a; Darley et al., 1969b). Hawthorne effects, when the patient is aware that their speech is being measured, may contribute to differences between clinical and environmental results (Deane et al., 2002). Communication partners of persons with PD complain that the patient's voice is soft and difficult to understand. These complaints may not match the patient's speech performance in a controlled clinical environment when the patient is focused on their speech performance. In controlled clinical environments, patients show the ability to adjust the intensity of their speech in clinical tasks and in conversation, yet overestimate the loudness of that speech compared with controls (Ho, Bradshaw, & Iansek, 2000; Miller et al., 2007). Questions have been raised in the literature as how clinical speech tasks generalize to the home environment (Keintz et al., 2007; Sarno, 1968; Weismer, 1984). Efficient and effective means of measuring speech in the home for baselines and treatment effects should be explored. The questions presented are: (1) What is the effect of environment on speech intensity and intelligibility? (2) What clinical tasks might predict intelligibility and intensity in the natural environment? (3) Are there treatment effects on spontaneous speech intelligibility and intensity within the home environment? And (4) What can we glean from our experience of measuring



speech in the home environment? These questions are answered and new methodologies for remote sensing and ambulatory measures of speech and voice are introduced in this project.

## **Significance**

This project is a first step in developing measures that could be used for evaluation of patients' speech functioning in their own natural environments. Research is needed to investigate whether or not there are discrepancies between clinical tasks and functional intelligibility in the natural environment, and the effect of treatment on functional intelligibility. Outcomes of this research may allow practitioners to gather data in the natural environment, for therapy planning and assessment purposes. As health care becomes more patient centered, the focus on a person's communication in their natural environment will be critical, and accurate assessment is needed. The World Health Organization states that health must include physical, mental and social state of being as well as environmental factors, melding into a multidimensional concept (World Health Organization, 2002). The aim of future research will be to explore new methods for the development of measures of speech and voice function in a patient's natural environment.

Objective measures of speech function in the natural environment remain illusive to speech-language pathologists. Natural environment measures of speech function may not correlate with clinical measures currently used for assessing speech or voice (Deane et al., 2002; Hunter, 2009; McAllister & Brandt, 2012). In patients with hypophonia, reports of quality of life are poorly related to acoustic measures (Wheeler et al., 2006). Perceptual ratings of speech by clinicians have poor inter-rater reliability and may not relate to objective acoustic measures of voice (Zraick et al., 2011). Therapeutic approaches for treatment of PD symptoms of hypophonia and dysarthria that could improve intelligibility have not been

measured in the natural environment for most well-known therapies (Deane et al., 2002).

### **Innovation**

LENA is a small wallet sized portable device that is encrypted and can store up to 16 hours of acoustic recordings. This device is referred to as a Digital Language Processor, or DLP. When worn by a patient, the DLP recordings include sound in the patient's environment and their speech, recorded in .wav format. Originally designed for recording environments of infants, software provides analysis of the language environment and measures of language development data such as mean length of utterance for children ages 2 to 48 months (Gilkerson & Richards, 2009). However, the recorded .wav file is easily extractable from the device for other types of voice and speech analysis. Praat speech acoustic software allows for accurate analysis of speech acoustics (Boersma & Weenink, 2010). As .wav files are the standard for acoustic analysis, these files can be played for listeners when transcribing speech content.

Ziaei, Sangwan and Hansen (2012) utilized LENA to monitor and model the natural/functional audio environment using healthy adults as their subjects. The researchers collected more than 35 recordings, each recording lasting 10 or more hours. This collection, which is known as the "Prof-Life-Log corpus," is yielding a unique and unprecedented opportunity to explore real-world natural audio samples. Ziaei et al. gathered two datasets from the corpus, (1) a controlled collection with homogenous recordings of various environments, and (2) a real-world naturalistic collection. Although this study was focused on modeling environmental/background noise, such an approach could be used to monitor spontaneous speech of patients with PD and healthy controls. Such measures could

then be related to patients' voice and speech measures in a clinical setting, and to develop pre- post- treatment measures. These will be the first steps to building an outcomes measurement system in the natural environment.

## **Approach**

**To determine the effect of the environment on speech intensity and intelligibility in persons with PD and healthy controls.**

**Problem.** Hypophonia accompanying PD is characterized by a soft voice and reduced intelligibility (Darley et al., 1969a; Darley et al., 1969b). Hawthorne effects may weaken the agreement between clinical and environmental results. Although patients are aware that their speech is being recorded in both settings, the focus on speech recording in the clinical setting when the patient is reading sentences and not formulating speech may differ from the natural environment. Listeners have difficulty understanding those with dysarthric speech (K. K. Tjaden & Liss, 1995), and these complaints may not match the patient's performance in a controlled clinical environment when the patient is focused on performing to their maximum capability (McCarney et al., 2007b). Others have suggested that patients with PD may report less severe intelligibility problems than their communication partners (Ho et al., 2000; Kalf et al., 2011; Miller et al., 2007).

**Purpose.** A study was developed to examine the effect of the environment on spontaneous speech.

**Hypotheses.** Speech intelligibility and intensity would differ significantly between clinical and natural environment settings, and the greater difference between environments would be found in persons with PD.

**Independent variable.** Environment (clinic, home).

**Dependent variables.** Spontaneous speech intelligibility and intensity.

## Methods.

At least 12 participants with PD would be recruited and participate. A power analysis was conducted based on pilot data from an undergraduate honor's project conducted by Lora Hellman in 2012, entitled, "Comparison of Speech Amplitude and Intelligibility Testing with Measures of Speech Communication in Daily Life in Parkinson Disease." Percent intelligibility in the clinical and natural environments was measured in 4 participants with PD and one healthy control and found effect size differences between the clinic and natural measures ranging from 0.38 to 1.47 for intelligibility and 0.41 to 1.8 for intensity in the four patients. SYSTAT software was used to compute the numbers of participants needed (Fig. 1) for a power of 0.80. For percent intelligibility, 11 subjects would be required to find differences on measures of speech intelligibility at  $p=0.025$  (Bonferroni corrected for 2 outcome variables) between clinic and natural environment in the patient's home. Further, for sound pressure level change scores, 8 subjects would be required to find significant differences at  $p=0.025$  (Fig. 2).

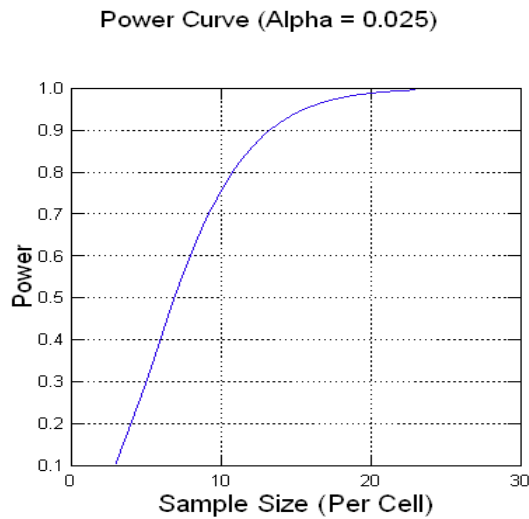
*Fig.1 Power analysis using pilot test data on 4 PD patients on difference in speech intelligibility*

Paired t-test with alternative 'not equal'

Expected difference	Standard Deviation of Difference	Effect Size	Alpha	Power	Non-Centrality Parameter
17.211	15.277	1.084	0.025	0.8	$1.084 \times \sqrt{\text{sample size}}$

Sample size (per cell)	7	8	9	10	11
Power	0.512	0.606	0.688	0.757	0.813

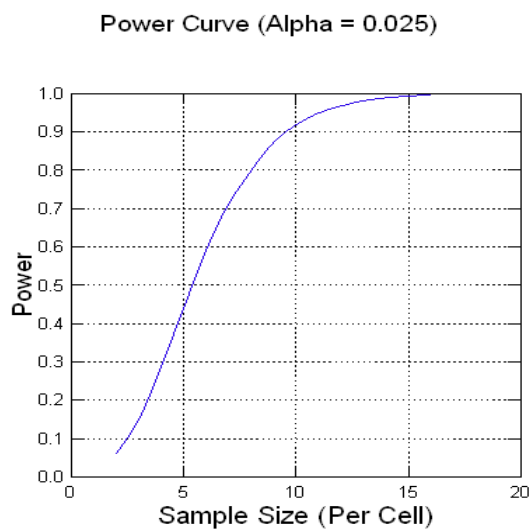


*Fig. 2 Using pilot test data on 4 PD patients on change in sound pressure level*

Paired t-test with alternative 'not equal'

Expected difference	Standard Deviation of Difference	Effect Size	Alpha	Power	Non-Centrality Parameter
5.698	4.226	1.348	0.025	0.8	$1.348 \times \text{SQRT}(\text{sample size})$

Sample size (per cell)	4	5	6	7	8
Power	0.286	0.443	0.591	0.713	0.806





Within group (clinic versus home environment) and between group (PD versus HV) comparisons and their interactions were completed to examine speech intelligibility and speech intensity (sound pressure level) in everyday life in 13 persons with PD and a control group of 12 healthy volunteers with self-reported normal speech. Differences between speech intelligibility and speech intensity (sound pressure level) in clinical testing versus natural environment and interactions between group and environment were examined.

Subjects: Volunteer participants were recruited from area PD support groups as well as via flyers and word of mouth in the community. Inclusion criteria included: over 50 years of age, a score of 23 or higher on the Folstein Mini Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975), diagnosis of PD by a neurologist, speech/voice impairment typical of hypophonia and/or dysarthria commensurate with PD as judged by a practicing speech-language pathologist with experience in neurological disorders, typical hearing abilities, as demonstrated by audiological screening at 40 dB at 500, 1000, 2000, 3000 and 4000 Hz (allowing for 4000 Hz notch if patient reports that they have experience shooting guns). Exclusion criteria include: history of voice or speech problems prior to PD diagnosis, diagnosis of other neurological diseases other than PD, current diagnosis of mental illness except depression, and non-native American English speaker. Healthy controls were recruited via word of mouth and JMU emails to faculty and staff. Inclusion criteria for healthy controls were: 50 years old or older, typical hearing abilities as demonstrated by an audiological screening at 40 dB at 500, 1000, 2000, 3000 and 4000 Hz (allowing for 4000 Hz notch if patient reports that they have

experience shooting guns), and an MMSE score of 23 or above. Exclusion criteria included: presence of neurological diseases, any speech impairments, current diagnosis of mental illness except depression, non-native speakers of American Standard English. All volunteers had telephone screening using the above criteria. The MMSE and hearing screening, were administered on the first visit after the consent process.

Tasks: All patients gave written consent in person prior to the start of the session. Patients completed a health questionnaire and the Voice Handicap Index (VHI) (Jacobson et al., 1997). Participants were then administered a hearing screening and the Folstein Mini Mental State Examination (MMSE) (Folstein et al., 1975). The following speech tasks were completed: conversation for a spontaneous speech sample (Appendix 1), the Sentence Intelligibility Test (Yorkston et al., 1996), and a scripted recording that includes maximum phonation time, maximum/minimum fundamental frequency, 4 different loudness levels on single words, repeated glottal stops and voiceless consonants between vowels, stressed words in sentences, intonation in sentences, and a counting task from 60 until the patient runs out of air (for maximum phonation time in speech with voiced and unvoiced sounds). The participants were then trained on the recording device.

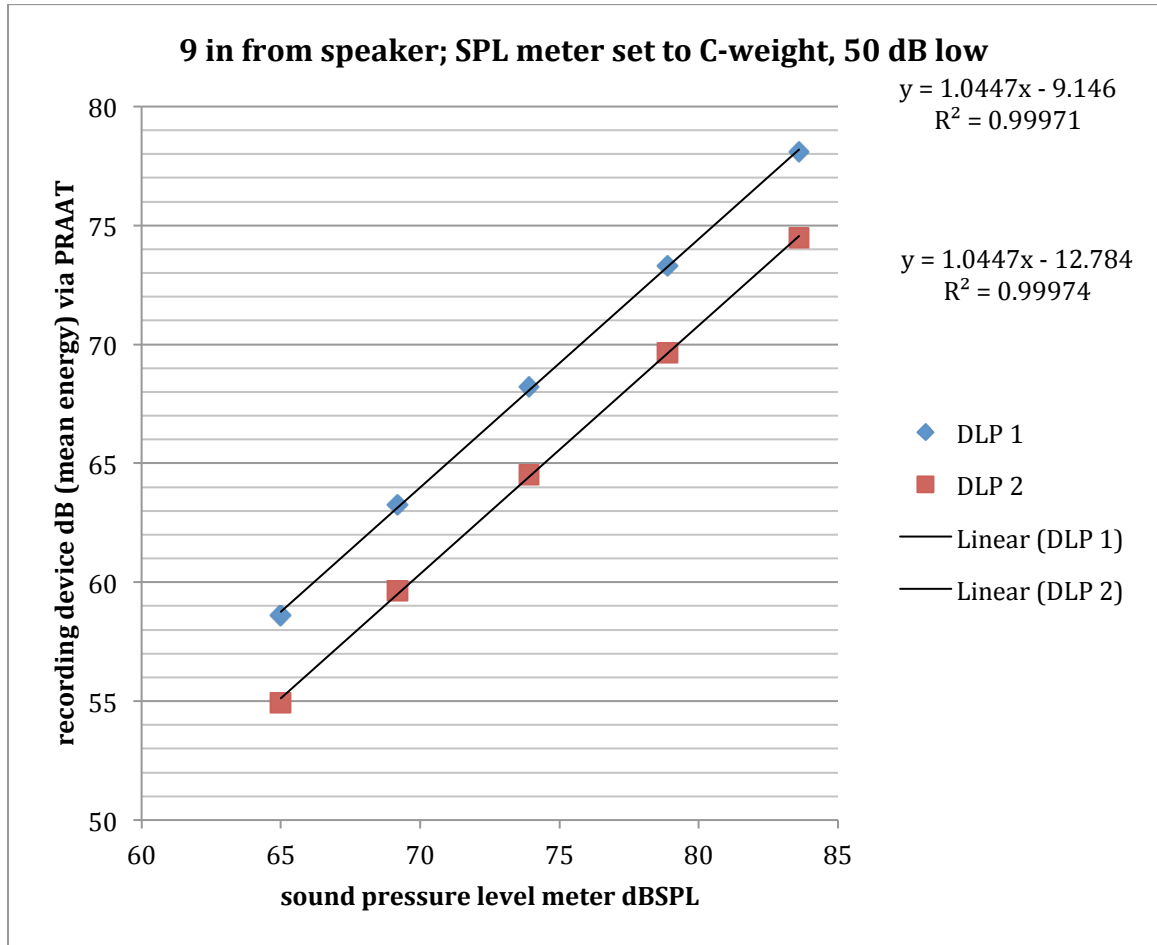
Instruments and settings: Clinical recordings were made in a sound attenuated room. A recording of the entire session was made with the LENA digital language processor (DLP). The DLP collects data using an omnidirectional microphone with a flat 20-20 kHz frequency response. Frequencies above 10 kHz are suppressed, as they are unlikely to contain human speech activity. Low

frequency data are suppressed through a 70 Hz high-pass filter. Anti-aliasing filtering is applied using 10 kHz low-pass filter to suppress high-frequency sounds prior to digitization using a 16 kHz 16-bit sigma-delta analog to digital (ADC) converter with 8x over-sampling digital interpolation. When the LENA DLP was tested with 500Hz calibration tones in a sound attenuated room, the hardware limiter for intensity suppression functionally activated at approximately 87dbSPL. This limiter was not expected to be problematic for the population of this study, given that normal conversational speech hovers around 60-70dSPL at 9 inches mic-to-mouth distance. When tested on various frequencies, the DLP showed response to 125 Hz, 250 Hz, 500 Hz, 750 Hz, 1000 Hz, 1500 Hz, 2000 Hz, 3000 Hz, 4000 Hz, 6000 Hz, and 8000 Hz. The DLP was placed on the participant at a 9-inch mouth-to-microphone distance, clipped on their chest just below their mouth, to record spontaneous speech and the assessment tasks. Natural environment recordings were made in the participants' home and other settings. Researchers took care to instruct patients on maintaining a microphone-to-mouth distance. They were provided with a ribbon to measure this distance to maintain the same mouth-to-microphone distance over the two days of recording. They were also provided with basic instructions for making recordings with the DLP.

Calibration: The DLPs were calibrated using 5 calibration tones (65, 70, 75, 80 and 85 dB SPL at 500 Hz). The tones were delivered via Grason-Stadler GSI 61 clinical audiometer and factory Grason-Stadler speakers in a sound attenuated room, to the DLP and an SPL meter (set at 50 dB low, C-weight, and fast speed) both 9 inches from the speaker. These calibration recordings were saved for analysis when

a linear interpolation was used to get the slope and y-intercept of the calibration recordings to be able to convert the dB into dB SPL. Calibration was recorded prior to each participant in the same fashion. The basic linear interpolation graph for the DLPs involved in this study is below.

*Fig. 3 Basic linear interpolation graph for calibrating the recording devices*



Measures and Data Analysis: For the purposes of this study, the following table demonstrates the data that was collected, the format for saving the data, measures, and analysis.

*Table 1: Data, format, measures and analysis for environment study*

Task	Equipment	Format	Measure	Analysis	Final Value
Clinical Spontaneous Speech Sentences	LENA DLP	.wav	Intensity	Mean intensity of utterance via PRAAT	<b>dB-SPL</b> processed via linear interpolation from calibration tone values
			Intelligibility	3 inexperienced listeners' transcriptions of sentences	Words correct/total no. of words=% <b>word intelligibility</b> in sentences
Natural Environment Spontaneous Speech Sentences	LENA DLP	.wav	Intensity	Mean intensity of utterance via PRAAT	<b>dB-SPL</b> processed via linear interpolation from calibration tone values
			Intelligibility	3 inexperienced listeners' transcriptions	Words correct/total words=% <b>word intelligibility</b> in sentences

Criteria for choosing spontaneous speech utterances in the clinical setting:

1. Participant's attention must not have been on the recording device for training or inquiry within 1 minute. For example, when reviewing the recordings to sample spontaneous speech in the clinic, one minute before and after any mention of the recording device, either for training proposes, placement, or spontaneous comments or questions about the device, is excluded.
2. Participant must have been engaged in casual conversation and not involved in an evaluative task. Any spontaneous utterances occurring during or between evaluative tasks were excluded from the clinic spontaneous speech sample that was analyzed.
3. Preferably, samples are taken during the interview portion at the beginning of the session, prior to being familiarized with any evaluative tasks. Patients give consent face to face prior to the session, so as soon as the session

starts, the LENA is placed on the patient, and the researchers say, “This is the recording device that we talked about and we will talk more about it at the end of the session.” Researchers then go on to engage the participant in conversation while they “just make a few notes,” so that the participant’s attention is not directed towards the recording device. Researchers were careful to initiate conversation about non-voice/speech topics, such as grandchildren, how far the patient had to drive, if they are planning to do anything while they are in town, etc. A list of potential topics to initiate is provided for the researchers (Appendix I). The intent is to match the natural environment as closely as possible with regard to conversational topics and participant behavior.

4. Sentences were chosen based on similar length and complexity to sentences in the Sentence Intelligibility Test (SIT) (Yorkston et al., 1996).

Spontaneous speech sentences were also drawn from the natural environment, representative of conversational speech over 2 days in a 16-hour sample (8 hours each day), using the following procedures and criteria:

1. Condense the 16-hour file by removing all non-speech sections, as well as clinical tasks.
2. Make note of how many minutes remain (for example 120 minutes)
3. Using a random number generator, generate a list of numbers representative of the number of minutes of speech in the reduced file (use [www.random.org](http://www.random.org))

4. Choose the minute timestamp with the first random number, mark with cursor and choose the closest sentence that the patient uttered from that marker. Make a note of the timestamps of the beginning and end of this sentence.

5. Move to the next random number and repeat until there are 12 sentences that meet the following criteria:

- Background noise is not louder than the target utterance.
- If background noise exists, it is consistent, for example, no phone ringing in the middle of utterance.
- Conversational partner does not interject in the utterance.
- Sentences as a group fit the profile of the sentence intelligibility test (4-16 words, use patient's SIT stimuli for reference), and are as similar as possible to the SIT sentences with regard to grammatical complexity and intonation.

6. Continue to repeat the process with the random number generator until 12 sentences are chosen that match closely to number of words and grammatical complexity of SIT stimuli for that patient

Intelligibility tasks were recorded using the DLP and transferred to a digital (.wav) file for analysis. Three different novel, inexperienced listeners transcribed each speaker's sentences, with no one listener hearing any repeated sentences. Each volunteer listener passed a hearing screening (at 30 dB at 500, 1000, 2000, 3000 and 4000 Hz) and was limited to 20 minutes of listening to reduce listener fatigue. They were provided with a laptop to transcribe what they hear as the sentences are played for them, or they used their own laptop. Prior to playing each

set of sentences, amplification levels on the speakers were set to match the same sound pressure level meter values read during the calibration tone session, by playing the calibration tones and reading the same sound level meter and adjusting the amplification level so that the sound level meter reading for the calibration tones match. Listeners were instructed that they would hear each sentence only once and that they are to transcribe exactly what they hear, without regard for punctuation. They are told not to make any corrections until the entire sentence is transcribed, to reduce the load on memory. An Intraclass Correlation Coefficient (ICC) revealed interrater reliability of .966, which exceeded our predetermined minimum of 0.65. Researchers transcribed the recordings according to the guidelines in the SIT manual that lent a “percentage of intelligibility.”

Comparisons between the listener transcriptions and the key made by the researchers were made with R, Package ‘qualV’ (van den Boogaart, Rost, Petzoldt, & Petzoldt, 2014) using the Least Common Sequence (LCS) function algorithm. The spreadsheet of all the responses was checked for spelling errors and then the LCS program was applied to the responses. Text was converted to lowercase without punctuation, erroneous spacing was removed, and the algorithm created output that examined for the number of identical words in the same sequence as the key. This number was divided by the greatest number in either sentence, accounting for errors of omission and commission. The mean of each set of 10 sentences was taken, multiplied by 100 and saved in the dataset. Reliability when compared to methods outlined in the SIT for percent intelligibility ranged from  $R^2=.98-1.0$ .

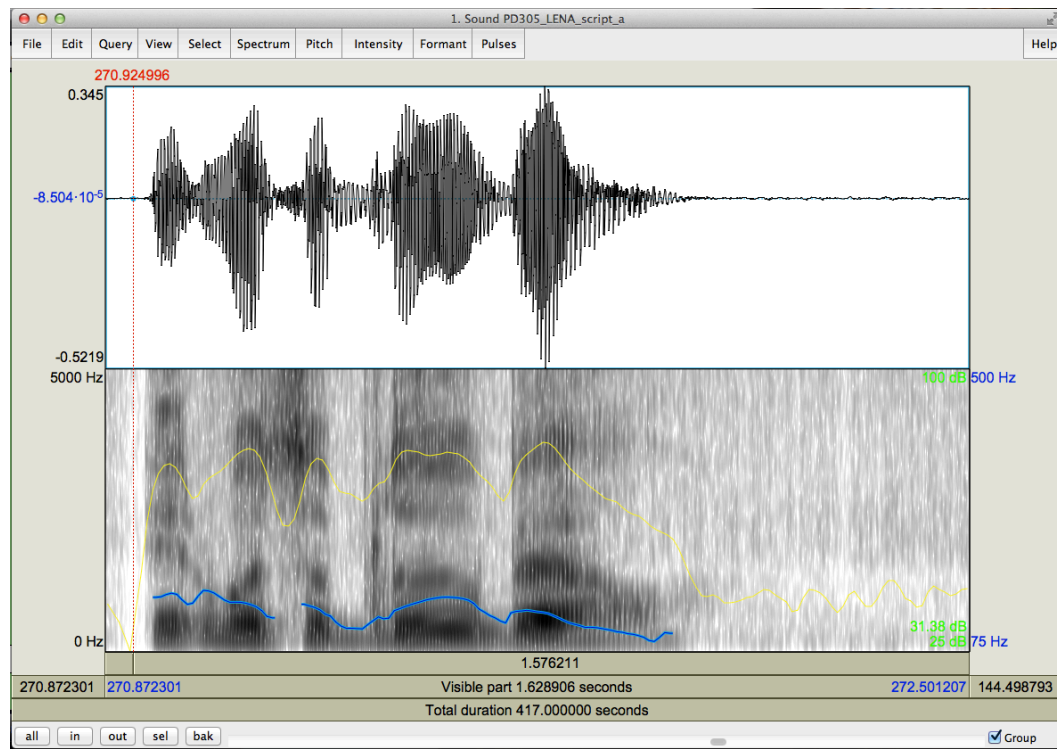


Intensity of the utterance was measured using Praat (Boersma & Weenink, 2010) software for speech analysis. Praat intensity settings were set to show a range of 25dB to 100dB and “subtract mean pressure” was checked as OFF. The vowel with peak amplitude in the sentence was selected and 5 cycles of that vowel in which the peak intensity level is most stable was chosen. The sentence was selected and if pauses exist greater than 150 ms, they were deleted. The mean intensity over the utterance was recorded. The intensity level was converted into decibels in sound pressure level re .0002 dynes/cm<sup>2</sup> using the slope and y-intercept from the equation derived by measurement of calibration tones derived from the same recording session for participant’s date, session and recording environment.

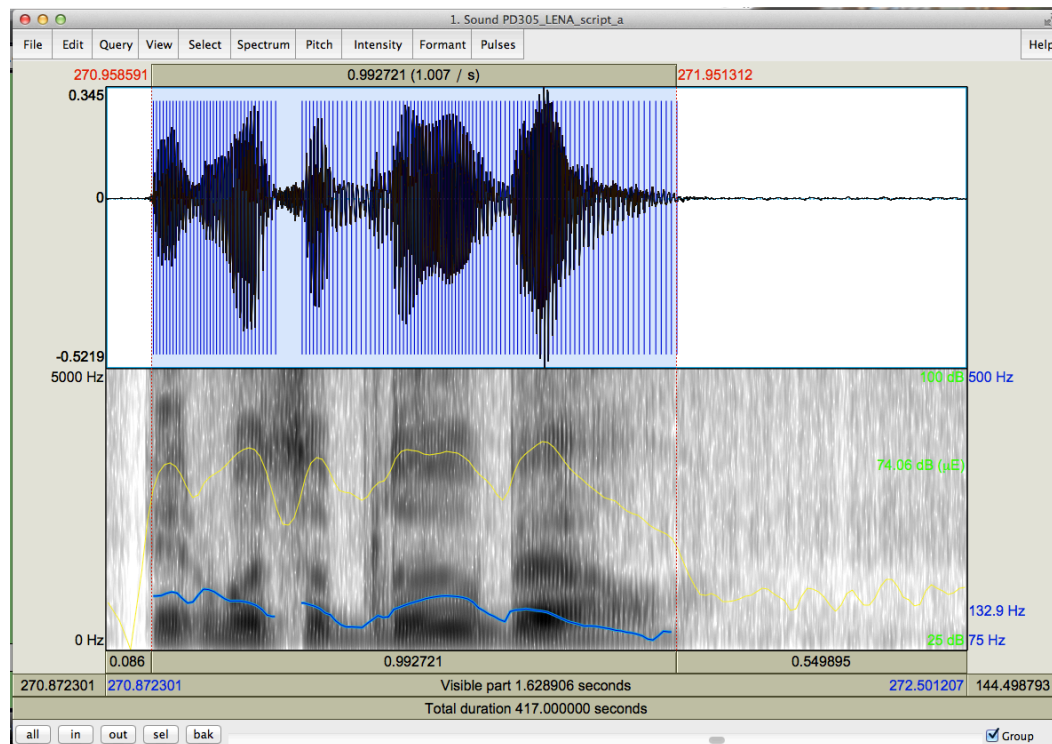
An example is listed below, “It was a blue bell.”

*Fig. 5 Analyzing intensity of a sentence via Praat*

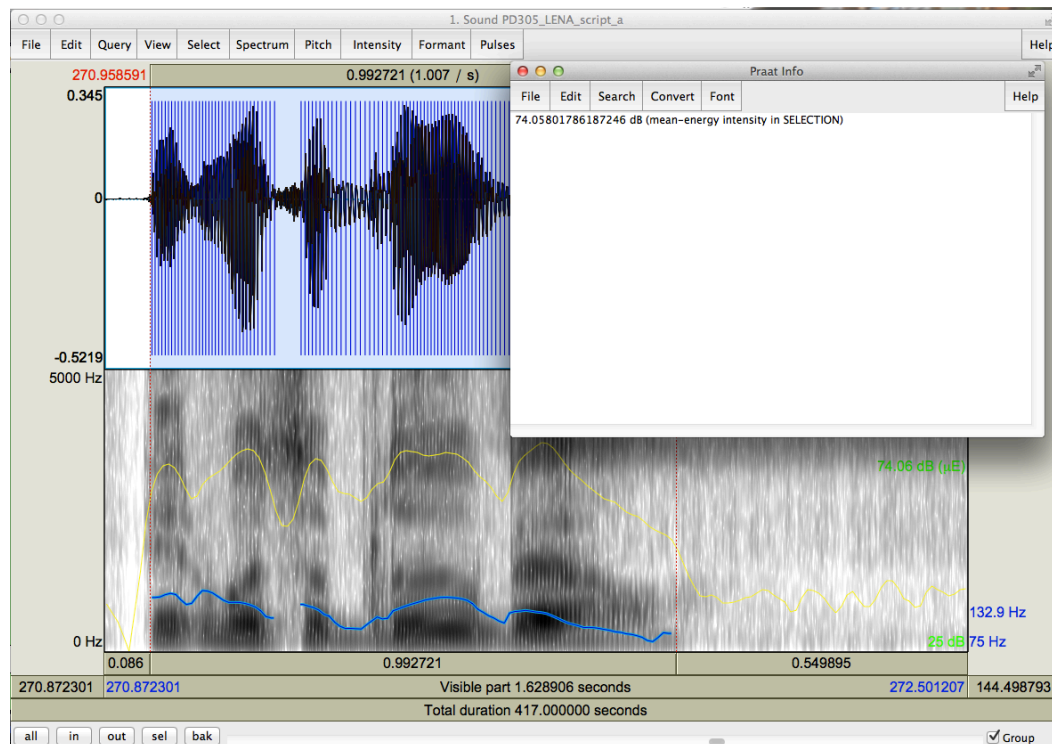
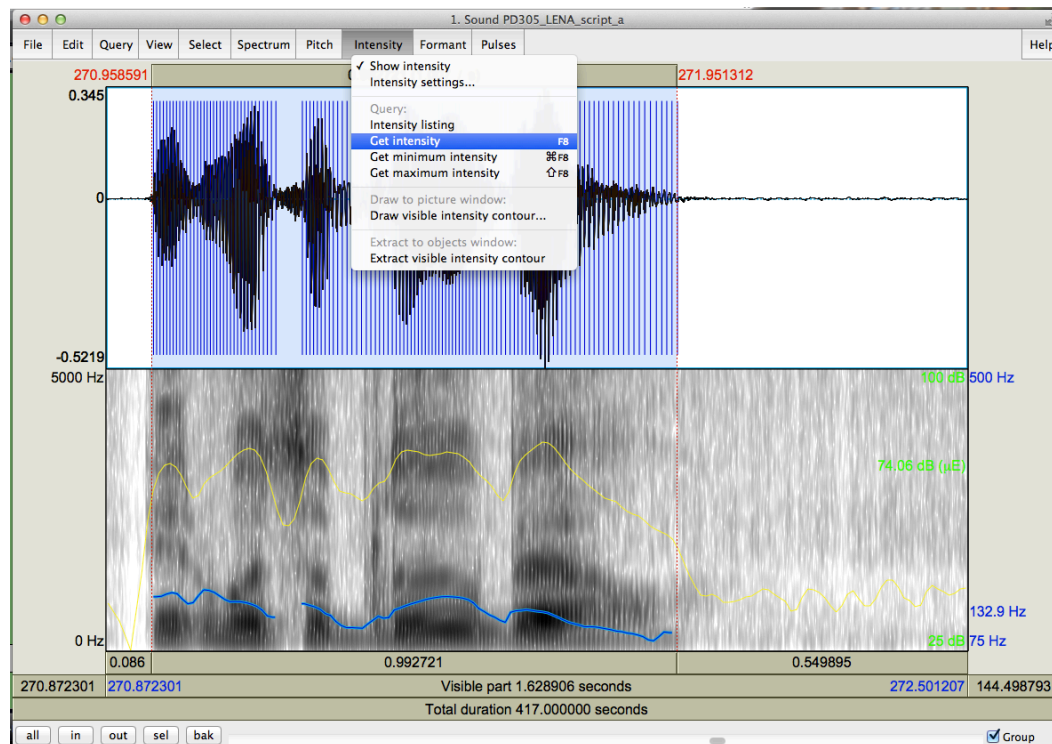
it wəz ə blu bəl



*Fig 5a. Select the sentence to be analyzed from the beginning to the end of the sound contour.*



*Fig. 5b. Select “intensity” and “get intensity,” or read intensity in green from the right side of the screen.*

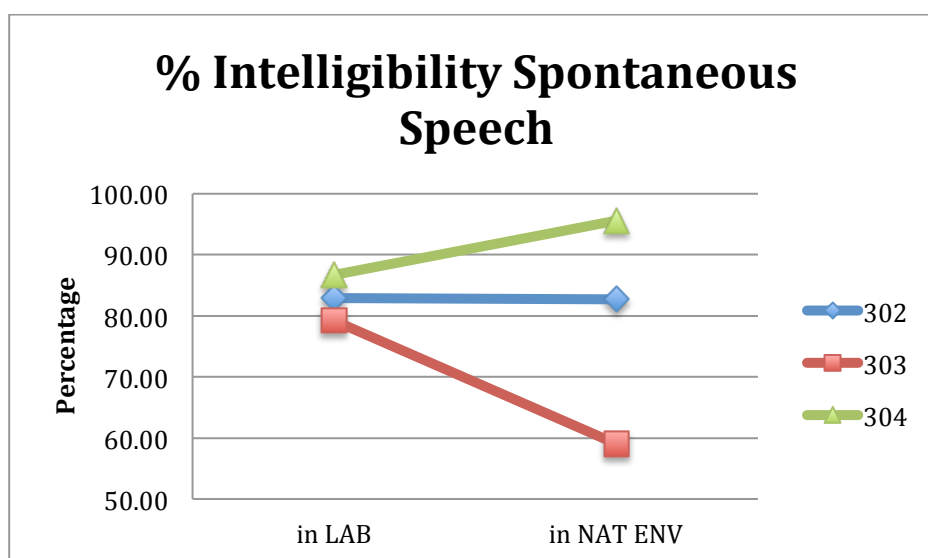
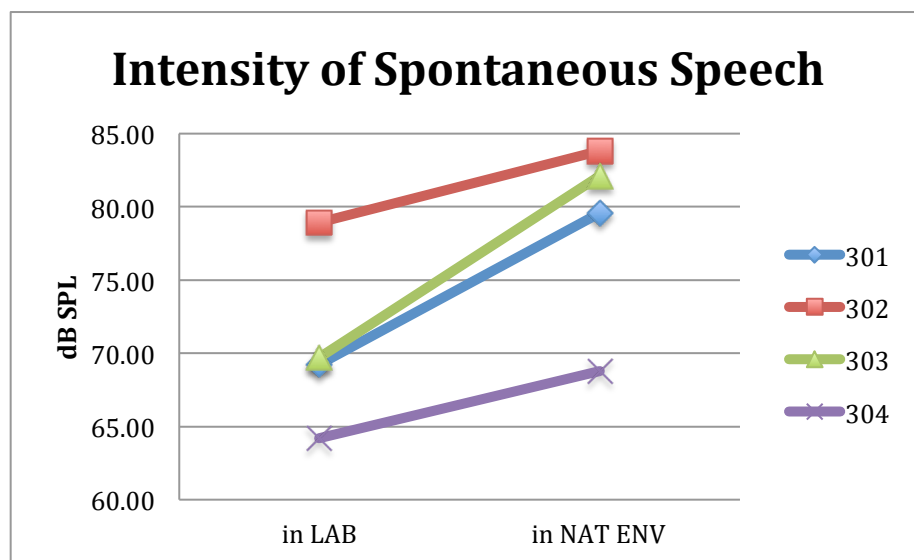


Record, in this case, mean energy from selection is 74.06 dB. This value was transformed via linear interpolation from calibration for that device and that day into dB SPL.

Statistical Analysis: Two-way repeated measures ANOVA compared clinic versus natural environment (environment) within subjects and PD versus healthy volunteers (group) between subjects and the interaction of environment and group. Separate ANOVAs were conducted for each dependent variable: spontaneous speech intelligibility, spontaneous speech intensity. Correlations were made to investigate the contributions of years since diagnosis, depression, background noise, and how intensity and intelligibility are related, as well as intelligibility between environments.

Pilot data: A pilot study was completed after 4 participants. All data below reflects spontaneous speech recorded via LENA DLP. Intensity data is in dB SPL (transformed from raw intensity readings via linear interpolation of SPL readings of calibration tones for that session). Intensity of speakers for intelligibility judges has also been adjusted to match SPL readings for the day of the patient session. Please note that the intensity pilot data below was taken from the stressed vowel in each sentence. The analysis method changed after the pilot study to the above method of mean intensity of the entire sentence, which we believe is more representative of connected speech and also will more easily show a relationship to intelligibility of the entire sentence.

*Fig. 6. Pilot data for environment study*



	Intensity		Intelligibility	
	in LAB	in NAT ENV	in LAB	in NAT ENV
301	69.23	79.60		68.10
302	78.93	83.80	82.91	82.66
303	69.70	82.10	79.22	59.00
304	64.20	68.80	86.68	95.60
mean	70.52	78.58	82.94	76.34

Pilot data showed that change in intelligibility measures between the clinical measures and the natural environment varies between patients.

Overall the SPL of speech increased in the natural environment from the clinic while no group trend was seen on percent intelligibility scores between the clinic and the natural environment.. Based on these three subjects intelligibility observed in the clinical environment may not relate to intelligibility in the natural environment. Also, it is worth noticing that patient 304, with the lowest intensity voice (in both environments) is also the MOST intelligible (in both environments). Conversely, patient 303, with one of the higher intensity voices in the lab, is the LEAST intelligible in the natural environment.

*Table 2: Coefficients of variation*

Clinical Intensity	Natural Environment Intensity	Clinical Intelligibility	Natural Environment Intelligibility
0.08704	0.08579	0.04497	0.21115

Coefficients of variation indicate greatest variability between subjects in the measure of intelligibility of spontaneous speech within the natural environment, though variability of intelligibility in the clinical environment was comparatively low.

Pilot data showed enough variance to move forward with the full study, which is reviewed in the Manuscript section of this document.

**To explore which clinical task predicts speech intelligibility in the natural environment.**

**Background.** Standard clinical tasks for evaluation of hypophonia and dysarthria reflect ASHA's preferred practice patterns for speech-language pathologists (SLPs) (American Speech-Language-Hearing Association, 2004) on voice and motor speech assessments. These include: mean sound pressure level (SPL) in prolonged vowel, range of SPL (SPL- $\Delta$ ), Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) (Kempster et al., 2009), Sentence Intelligibility Test (SIT) (Yorkston et al., 1996), maximum phonation time (MPT), harmonic to noise ratio (HNR), and range of  $F_0$  ( $F_0$ - $\Delta$ ). Clinicians will also often include quality of life scales such as the Voice Handicap Index (VHI) (Jacobson et al., 1997; C. A. Rosen, Lee, Osborne, Zullo, & Murry, 2004). The effects of task have been demonstrated within the clinical environment (Ho, Iansek, & Bradshaw, 1999; Huber & Darling, 2011). Clinical tasks have also been related to clinical ratings of speech clarity (Kim, Kent, & Weismer, 2011). To determine if previous studies have examined the relationship between clinical measures and measure in the natural environment a search was conducted on the PubMed and Scopus/Elsevier databases using the following terms: "acoustic, speech, measures, conversation." PubMed revealed 23 articles, with 3 mentioning PD, one of which was regarding telerehabilitation. Of the remaining two, the difference between acoustic measures of hypokinetic dysarthric conversational speech and conversational speech of healthy controls was investigated (K. M. Rosen & Duffy, 2006), as well as the effect of task on volume (Ho et al., 1999). Scopus/Elsevier revealed 77, with 3 having "Parkinson" in their titles,



one of which was regarding telerehabilitation. Of the remaining two, one examined the effect of task on dysfluencies in PD speech, finding that dysfluencies were more abundant in conversation than in other speech tasks (Van Lancker Sidtis, Cameron, & Sidtis, 2012). The other examined the effects of pallidotomy surgery on sentence measures, also showing some differences in intelligibility between reading, picture description, and conversation (G. Schulz, 2004).

**Problem.** The clinical environment encourages patients to attend to *how* they are talking, whereas spontaneous speech in the natural environment requires patients to attend to *what* they are saying and not how they are talking (Connor & Abbs, 1991; Kempler & Lancker, 2002). Patients' family members often complain of problems with hearing and understanding their speech in their natural environment; however, assessments take place in the clinical environment. Thus task effect differences may play a role (Kempler & Lancker, 2002; R. Kent et al., 1989) as well as combination of speech and non-speech tasks, such as what is seen during conversational speech in the natural environment (Bunton & Keintz, 2008). In the study on environment (detailed in Manuscript section), we saw differences between environments that were unexpected, such as increased speech intensity in the home environment, and intelligibility unaffected by environment. Persons with PD were set apart by decreased intelligibility compared to healthy controls, even though the sample consisted of mild to moderately impaired persons. Initial correlations between years since diagnosis, depression, and background noise with intensity and intelligibility in the home were not significant. Only comparisons between clinic spontaneous speech intelligibility and home spontaneous speech

intelligibility were significant. However, there is not a standardized way to look at spontaneous speech intelligibility in the clinic, therefore predictions can not be made to spontaneous speech intelligibility in the home, the variable that sets apart the participants with PD. Clinical tasks need to be explored to see if any typical clinical task(s) might be predictive of spontaneous speech in the home.

**Purpose.** To explore the relationship of standard clinical measures to intelligibility within the natural environment. These measures included mean sound pressure level in prolonged vowel, Consensus Auditory Perceptual Evaluation of Voice (CAPE-V), Sentence Intelligibility Test (SIT), maximum phonation time, harmonic to noise ratio, range of  $F_0$ . Improved intelligibility in conversational speech should be the primary aim for evaluation and treatment, and likely differs from the usual clinical measures of voice (Deane et al., 2002).

**Outcome variable.** Spontaneous speech intelligibility in the natural environment ( $SI_{HOME}$ ).

**Predictor variables.** Originally, the following were selected for inclusion in our predictor variable set: mean sound pressure level ( $SPL_{VOWEL}$ ) in prolonged vowel, range of SPL ( $SPL-\Delta$ ), score on the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V), percent intelligibility based on the Sentence Intelligibility Test as analyzed via R ( $SIT_R$ ), maximum phonation time (MPT), harmonic to noise ratio (HNR), range of  $F_0$  ( $F_0-\Delta$ ). After finding that  $SPL-\Delta$  and  $SPL_{VOWEL}$  covaried ( $R^2=.651$ ),  $SPL-\Delta$  was eliminated from our dataset. Additional inquiry into best practices for range of  $F_0$  led us to use the interquartile range of  $F_0$  in targeted sentences ( $F_0-IQR$ )

(Baken & Orlikoff, 2000; Busso, Lee, & Narayanan, 2009). Our final set of predictors included:  $SIT_R$ , CAPE-V,  $SPL_{VOWEL}$ , HNR, MPT and  $F_0$ -IQR

### **Methods.**

Subjects: Data was taken from the PD group in the aforementioned environmental study.

Tasks: All patients gave consent in person prior to the start of the session. Patients also filled out a health questionnaire and the Voice Handicap Index (VHI) (Jacobson et al., 1997). Participants were administered the following in the clinical environment: hearing screening, Folstein Mini Mental State Examination (MMSE) (Folstein et al., 1975), engaged in conversation for spontaneous speech sample (Appendix 1), took the Sentence Intelligibility Test (Yorkston et al., 1996), and followed a script recording that included tasks for measures of fundamental frequency range, voice quality measures such as the harmonic to noise ratio (Yumoto, Gould, & Baer, 1982), maximum phonation time, measures of the intensity dB SPL range when practicing 4 different loudness levels in speech, repeated glottal stops between vowels and repeated voiceless consonants between vowels, stressed words in sentences, intonation within sentences, and a counting task from 60 until the patient runs out of air (for maximum phonation time in speech with voiced and unvoiced sounds).

Instrumentation: Clinical recordings were made in a sound attenuated room. A recording of the entire session was made with the LENA digital language processor (DLP). The DLP settings remained the same as previously described.

Calibration: As previously described, the DLP was calibrated using 5 calibration tones, delivered via Grason-Stadler GSI 61 clinical audiometer and factory Grason-Stadler speakers in a sound attenuated room, to the DLP and an SPL meter (set at 50 dB low, C-weight, and fast speed) distanced 9 inches from the speaker. These calibration recordings were saved for analysis when a linear interpolation was used to get the slope and y-intercept of the calibration recordings to be able to convert the dB into dB SPL. Calibration was ensured prior to each participant in the same fashion.

Measures and Data Analysis: For the purposes of this exploration, the following table lists the data collected (outcome variable is listed first), the format for saving the data, measures, and analysis. Both perceptual and acoustic measures were recorded using the LENA DLP for consistency.

*Table 3: Data, format, measures and analysis for exploratory study*

Task	Equipment	Format	Measure	Analysis	Final Value
Spontaneous Speech Sentences	LENA DLP	.wav	Intelligibility	3 inexperienced listeners' transcriptions	Words correct/total words=% <b>word intelligibility</b> in sentences for $SI_{HOME}$ , $SI_{CLINIC}$
Say "ah" until I say stop	LENA DLP	.wav	Intensity	Mean intensity of vowel via Praat	<b>dB-SPL</b> processed via linear interpolation from caltone values for $SPL_{VOWEL}$
Sentence reading	LENA DLP	.wav	(SIT)	3 inexperienced listeners' transcriptions	Words correct/total words=% <b>word intelligibility</b> in sentences for $SIT_R$
Clinical tasks	LENA DLP	.wav	(CAPE-V)	Via CAPE-V protocol	<b>Overall severity</b> score on a scale of 0-100 for CAPE-V.
Say "ah" for as long as you can	LENA DLP	.wav	Max phonation time	Praat > select from first to last cycle > voice report > time range of selection	<b>Number of seconds</b> to the first decimal for MPT.

Sustained “ah”	LENA DLP	.wav	Harmonic to noise ratio	Praat > select middle 90% of vowel > voice report > mean autocorrelation of harmonicity	Harmonicity expressed in dB, example: if .98 of the signal is periodic, HNR is $10 \cdot \log_{10}(98/1) = 19.9$ dB
Repeated sentences targeting intonation	LENA DLP	.wav	range of $F_0$	Praat > select from first to last cycle > voice report > mean and SD of $F_0$	Calculate interquartile range via standard z score, mean and SD for $F_0$ -IQR

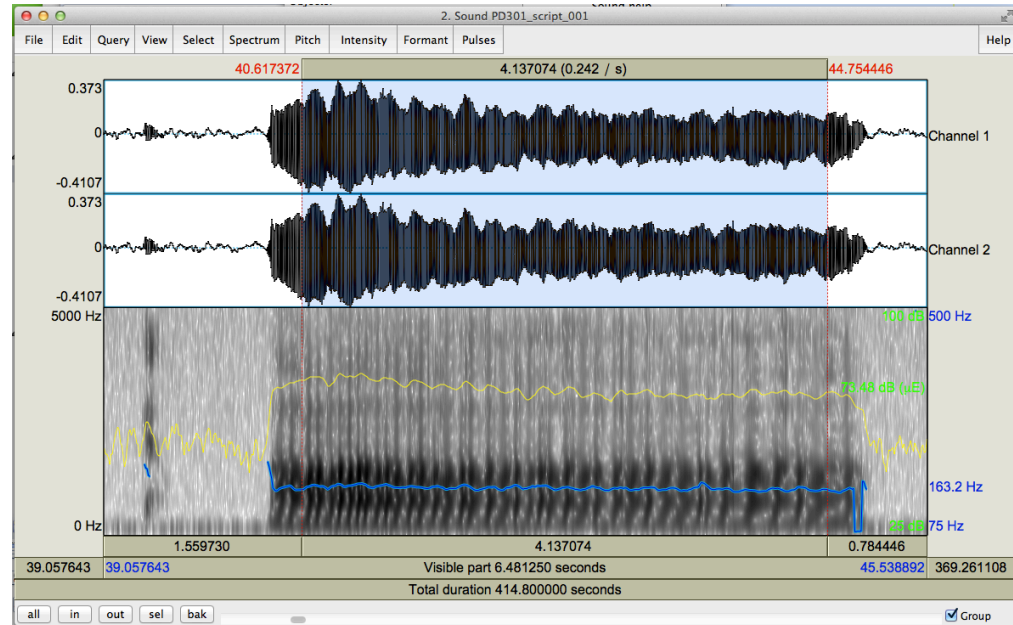
- Intelligibility measurements ( $SI_{HOME}$ ,  $SI_{CLINIC}$ ,  $SIT_R$ ): Natural environment

spontaneous speech sentences were randomly selected from the natural environment, representative of conversational speech over 2 days in a 15-hour sample (7-8 hours each day), using the same procedures and criteria as the environment study. SIT sentences were chosen via the SIT computer program (Yorkston et al., 1996) Three novel listeners transcribed each sentence using laptops and transcribing into a Microsoft Excel file.

Amplification levels on speakers matched the day of the recordings as checked by sound pressure level meter. Comparisons of transcriptions were made to a key that the researchers made (who were allowed to listen at higher levels if needed and listen for context, an unlimited number of times). The listener transcriptions and the key were compared via R, Package ‘qualV’ using the Least Common Sequence (LCS) algorithm (van den Boogaart et al., 2014). Therefore, we indicated ‘ $SIT_R$ ’ for this variable as distinct from ‘SIT’ for the commercially available assessment.
- Mean sound pressure level ( $SPL_{VOWEL}$ ) prolonged vowel: The middle 90% of the prolonged vowel task was selected, mean intensity reading was recorded from Praat (e.g. Fig.7), and transformed via linear interpolation from

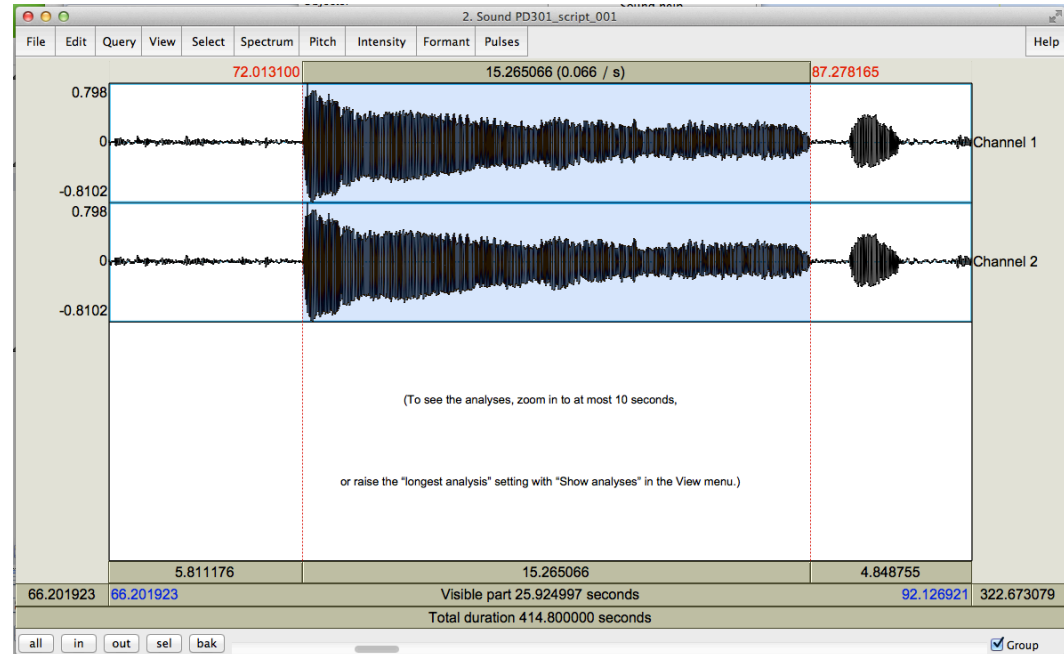
calibration recording. In the example below, Praat reads 73.48 dB, and this value was transformed for the final value via the slope and y-intercept from the calibration exercise for LENA microphone at 9 inches from the mouth.

*Fig. 7: Measuring  $SPL_{VOWEL}$*



- Consensus Auditory Perceptual Evaluation of Voice (CAPE-V): Using the CAPE-V standard protocol (Kempster et al., 2009) a speech-language pathologist with training on this method determined the overall score out of 100. Expert clinician was given a digital recording with 1-2 minutes of conversation, SIT sentences, and the script recording of speech tasks.
- Maximum phonation time (MPT): Using Praat (Fig.9), from first to last cycle of MPT task was selected, then voice report, number of seconds to the first decimal place were recorded. In this example, the value was 15.3 seconds.

*Figure 8: Measuring maximum phonation time*



- Harmonic to noise ratio (HNR): Using Praat, pulses were selected, then “show pulses” and all pulses of vowel from sustained “ah” were selected, then voice report, and mean harmonics-to-noise ratio was recorded. HNR was expressed as dB, for example: if .98 of the signal is periodic, HNR is  $10 \cdot \log_{10}(98/1) = 19.9$  dB
- Range of  $F_0$  ( $F_0$ -IQR): Using PRAAT the entire sentence was selected, then the mean pitch and standard deviation from each of the two sentence tasks was recorded. Then using standard z scores, the mean and the standard deviation, interquartile range was calculated, and an average over the two tasks was taken.

Statistical Analysis: multiple regression. Outcome variable:  $SI_{HOME}$ . Predictor variables:  $SIT_R$ , CAPE-V,  $SPL_{VOWEL}$ , HNR, MPT and  $F_0$ -IQR.

Results are discussed in the manuscript portion of this document.



**To assess the effect of treatment on speech intelligibility and intensity in the natural environment.**

**Background.** In seeking a treatment protocol in which to apply pre-post measures within the natural environment via LENA technology, the SpeechVive™ stood out as a treatment that exists within the natural environment that potentially functionally impacts intensity and intelligibility. Some literature has indicated sensory deficits in PD as related to hypophonia (Fox et al., 2002; Hammer & Barlow, 2010), while others point to coordination across neurological subsystems for motor execution (Stelmach et al., 1989). In the limb, fine motor movement, and gait literature, external cues have been shown to benefit patients both from a sensory and coordination perspective (Morris et al., 1994; Oliveira et al., 1997). Similar principles may also affect respiration and voice (Sadagopan & Huber, 2007). Progressive reductions in articulatory movement occur throughout the disease process in PD (Walsh & Smith, 2012), which may make PD patients more reliant on external cues for feedback. The SpeechVive™ is a wearable device that delivers cocktail chatter to one ear and thus elicits the Lombard Effect to speak more loudly due to the addition of background noise (Pick & al., 1989). As previously discussed, global treatment approaches to dysarthria in PD (i.e. rate reduction, increased loudness, and clear speech) are those that extend across an entire utterance and impact multiple speech components (i.e. respiration, phonation, articulation, resonance) (Yorkston et al., 2007). These treatment techniques are intended to improve intelligibility in dysarthria (Duffy, 2012; Fox et al., 2002; Weismer et al., 2012; Yorkston et al., 2007) by way of improving the multiple speech components,

including phonation. Instruction to speak louder has resulted in increase jaw movement (Connor et al., 1989; Connor & Abbs, 1991) and increased respiration (Huber & Darling, 2011) in patients with PD. Principles of practice and exercise physiology indicate that if someone practices heavier loading on the muscles he/she will become stronger (Muller, 1970; Taaffe et al., 1999). Currently Huber and colleagues are evaluating whether 12 weeks use of the SpeechVive improves speech intensity, intelligibility and other characteristics within the context of a multiple single subject design. Speech is tested with and without the device, in the presence of a speech pathologist, with a clinical setup within the home.

**Problem.** Pre- and post- test measures of spontaneous speech intelligibility and intensity in the natural environment are not included within the current SpeechVive™ study. Huber and colleagues seek natural environment measures to expand on the results from their clinical measures.

**Purpose.** To assess pre- and post- treatment measures of speech intelligibility and intensity, within the natural environment, in tandem with the SpeechVive™ study, particularly carryover effects. To examine the advantages and disadvantages of speech measurement in the natural environment for the purposes of measuring progress.

**Hypothesis.** Application of the SpeechVive™ device for 12 weeks will show an increase in intelligibility and intensity in the natural environment.

**Independent Factor.** Treatment

**Dependent variables.** Spontaneous speech intelligibility, spontaneous speech intensity.

**Methods.**

Subjects: Subjects were recruited from the environment study, and 7 of the 13 completed this treatment study.

Tasks: This study ran in tandem with the first study so the methodology is the same with regards to sampling criteria and data collection with a few additions: (a) participant will begin the SpeechVive™ study no longer than a week after completing the environment study, (b) participants will begin the treatment study within 24 hours after wearing the device for 12 weeks, per the study, that is the participants will end the SpeechVive™ study and begin this study within 24 hours.

Instrumentation: As with the environment study, a recording of the entire session was made with the LENA digital language processor (DLP). Refer to previous information on specifications.

Calibration: As was previously, the DLP was calibrated using 5 calibration tones, in a sound attenuated room, to the DLP and an SPL meter distanced 9 inches from the speaker. These calibration recordings were saved for analysis when a linear interpolation was used to get the slope and y-intercept of the calibration recordings to convert the dB into dB SPL. Calibration was ensured prior to each participant in the same fashion.

Measures and Data Analysis: For the purposes of this study, table 4 demonstrates the data that was collected (outcome variable is listed first), the format for saving the data, measures, and analysis.

*Table 4: Data, format, measures and analysis for treatment study*

Task	Equip- ment	For- mat	Measure	Analysis	Final Value
Natural Environment Spontaneous Speech Sentences	LENA DLP	.wav	Intelli- gibility	3 inexperienced listeners' transcript- tions	Words correct/total words=% <b>word intelligibility</b> in sentences via R
Natural Environment Spontaneous Speech Sentences	LENA DLP	.wav	Intensity	Mean intensity of utterance via PRAAT	<b>dB-SPL</b> processed via linear interpolation from caltone values

- Intelligibility ( $SI_{HOME}$ ): Natural environment spontaneous speech sentences was drawn from the natural environment, representative of conversational speech over 2 days in a 16-hour sample (8 hours each day), using the same procedures and criteria as in the environment study.
- Sound pressure level of the sentence ( $SPL_{SENT}$ ): In similar fashion to mean intensity reading the environment study, extract pauses > 150ms, select entire utterance, record mean intensity reading from PRAAT, and transform via linear interpolation from the calibration recording.

Statistical Analysis: Repeated measures ANOVA will compare pre- and post-treatment values of intelligibility and intensity in spontaneous speech within the natural environment.

Pilot Data for a power analysis: Preliminary data on Intelligibility was not available. Pre- and post- measures on dB SPL were available on only 3 patients.

Note: intensity values are from stressed vowel in each sentence, rather than the full sentence.

*Table 5: Pilot data for treatment study*

	Intensity	
	PRE	POST
301	75.9	79.6
302	83.8	84.2
307	88.1	88.7

The effect size for difference in dB pre and post is 0.84. Using G power to do a power analysis the results of the required N for aim 3 are 11 subjects (see below). Seven subjects completed the study. Since our pre-determined n was not achieved, we wrote this up as an expanded pilot study for feasibility. The results are discussed in the pilot study section of this document.

#### **G Power 3.1 output**

**t tests** – Means: Difference between two dependent means (matched pairs)

**Analysis:** A priori: Compute required sample size

**Input:** Tail(s) = One

Effect size dz = 0.84

$\alpha$  err prob = 0.05

Power (1- $\beta$  err prob) = 0.80

**Output:** Noncentrality parameter  $\delta$  = 2.7859648

Critical t = 1.8124611

Df = 10

Total sample size = 11

Actual power = 0.8278205

#### **Potential Limitations**

Inherent in the natural environment is variability. Some variables cannot be controlled for, such as number and quality of communication partners, amount of interaction, speaking environments (home, church, neighbor's home, etc.) cues within the natural environment that may enhance behaviors, or items in the natural

environment that may inhibit certain behaviors. Several factors influence the voice and speech production in persons with PD such as motor execution deficits, task effects, cognition and linguistic processing demands. In addition disease factors and ongoing neuropharmacological treatment can alter speech, as shown in what is our Fig. 9, below, but Fig. 1 in the original article (Goberman & Coelho, 2002).

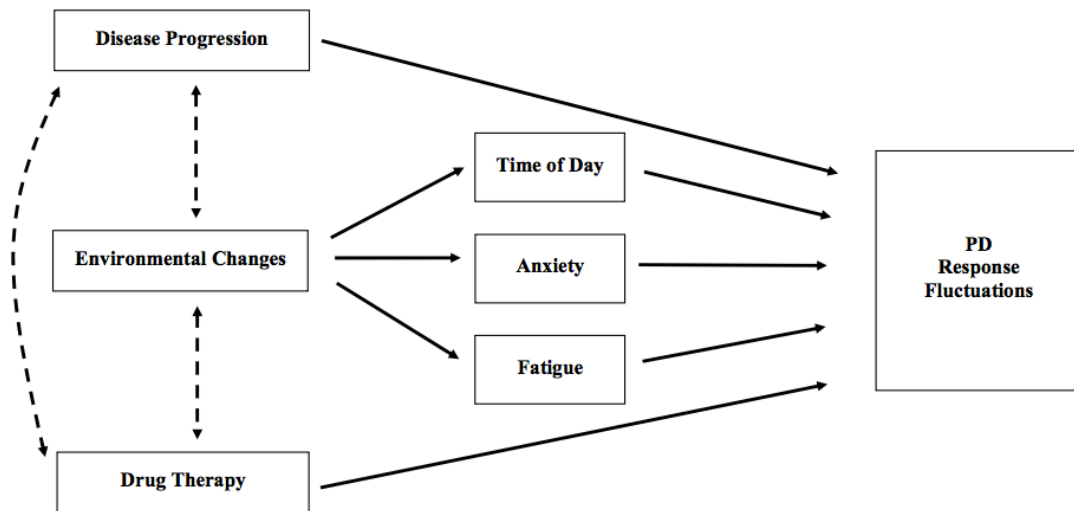


Fig. 1. Factors contributing to response fluctuations in PD. Dotted arrows represent the relationship between disease progression, environmental changes, and drug therapy changes in determining the response fluctuations. Solid arrows represent factors contributing to response fluctuations.

Reductions in variance may occur when exploring the same natural environment, within subjects, such as in the Pilot study (pre- post-treatment). It is expected that the environments themselves vary; therefore differences between groups may be more variable than individual differences pre- and post- within subjects. However, this study may be limited in number of participants according to device availability for the study.

Intelligibility may depend on other variables such as test material, personnel, training, test procedures, and state of the speaker (R. Kent et al., 1989). The test material and procedures are controlled in clinical assessment but they may be less relevant in the natural environment. The use of speech for communication within

the natural environment, in terms of functional intelligibility, is our variable of interest.

Other limitations relevant in the natural environment is the time that medication is taken relative to the recording, the patient's feeling of fatigue and level of distraction by preceding or concurrent events. Such "limitations" contribute to the effect of the natural environment, within and between subjects. Limitations are more fully discussed in the manuscript and pilot.

## References

- American Speech-Language-Hearing Association. (2004). *Preferred practice patterns for the profession of speech-language pathology [preferred practice patterns]*. ().
- Asgari, M., & Shafran, I. (2010). Predicting severity of parkinson's disease from speech. Paper presented at the *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 5201-5204.
- Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* Cengage Learning.
- Baker, K. K., Ramig, L. O., Luschei, E. S., & Smith, M. E. (1998). Thyroarytenoid muscle activity associated with hypophonia in parkinson disease and aging. *Neurology*, 51(6), 1592-1598.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio*,
- Boersma, P., & Weenink, D. (2010). {P} raat: Doing phonetics by computer.
- Boersman, P., & Weenink, D. *Praat: Doing phonetics by computer [computer program]*
- Braak, H., Tredici, K. D., Rüb, U., de Vos, R. A., Jansen Steur, E. N., & Braak, E. (2003). Staging of brain pathology related to sporadic parkinson's disease. *Neurobiology of Aging*, 24(2), 197-211.
- Buder, E. H., & Cannito, M. P. (2009). Acoustic assessment of vocal function. *Neurologic disorders of the larynx* (pp. 74-84) Thieme Medical Publishers New York.



- Bunton, K., & Keintz, C. (2006). Effects of a concurrent motor task on speech intelligibility for speakers with parkinson disease. Paper presented at the *Conference on Motor Speech*, Austin, TX,
- Bunton, K., & Keintz, C. K. (2008). The use of a dual-task paradigm for assessing speech intelligibility in clients with parkinson disease. *Journal of Medical Speech-Language Pathology*, 16(3), 141-155.
- Burn, D. (2000). Parkinsons disease: An overview. *The Pharmaceutical Journal*, 246, 333-337.
- Busso, C., Lee, S., & Narayanan, S. (2009). Analysis of emotionally salient aspects of fundamental frequency for emotion detection. *Audio, Speech, and Language Processing, IEEE Transactions On*, 17(4), 582-596.
- Connor, N. P., & Abbs, J. H. (1991). Task-dependent variations in parkinsonian motor impairments. *Brain : A Journal of Neurology*, 114 ( Pt 1A)(Pt 1A), 321-332.
- Connor, N. P., Abbs, J. H., Cole, K. J., & Gracco, V. L. (1989). Parkinsonian deficits in serial multiarticulate movements for speech. *Brain*, 112(4), 997-1009.
- Countryman, S., & Ramig, L. (1993). Effects of intensive voice therapy on voice deficits associated with bilateral thalamotomy in parkinson's disease: A case study. *Journal of Medical Speech-Language Pathology*, 1(4), 233-250.
- Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population-based norms for the mini-mental state examination by age and educational level. *Jama*, 269(18), 2386-2391.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research*, 12, 246-269.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Clusters of deviant speech dimensions in the dysarthria. *Journal of Speech and Hearing Research*, 12, 462-496.

De Bodt, M. S., Wuyts, F. L., Van de Heyning, Paul H, & Croux, C. (1997). Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11(1), 74-80.

de Lau, L. M., & Breteler, M. (2006). Epidemiology of parkinson's disease. *The Lancet Neurology*, 5(6), 525-535.

Deane, K. H., Ellis-Hill, C., Jones, D., Whurr, R., Ben-Shlomo, Y., Playford, E. D., & Clarke, C. E. (2002). Systematic review of paramedical therapies for parkinson's disease. *Movement Disorders*, 17(5), 984-991.

Del Tredici, K., Rub, U., de Vos, R. A., Bohl, J. R., & Braak, H. (2002). Where does parkinson disease pathology begin in the brain: *Journal of Neuropathology & Experimental Neurology*, 61(5), 413-426.

DeLong, M. R., & Georgopoulos, A. P. (2011). Motor functions of the basal ganglia. *Comprehensive Physiology*,

Dorsey, E. R., Constantinescu, R., Thompson, J. P., Biglan, K. M., Holloway, R. G., Kieburtz, K., . . . Tanner, C. M. (2007). Projected number of people with parkinson disease in the most populous nations, 2005 through 2030. *Neurology*, 68(5), 384-386.  
doi:01.wnl.0000247740.47667.03 [pii]

Dromey, C., & Benson, A. (2003). Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research*, 46(5), 1234-1246.

- Dromey, C., Ramig, L. O., & Johnson, A. B. (1995). Phonatory and articulatory changes associated with increased vocal intensity in parkinson DiseaseA case study. *Journal of Speech, Language, and Hearing Research*, 38(4), 751-764.
- Duffy, J. R. (2012). *Motor speech disorders: Substrates, differential diagnosis, and management* Elsevier Health Sciences.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G\* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189-198.
- Ford, M., Baer, C. T., Xu, D., Yapanel, U., & Gray, S. (2008). *The LENATM Language Environment Analysis System: Audio Specifications of the DLP-0121*,
- Forrest, K., Weismer, G., & Turner, G. S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults. *The Journal of the Acoustical Society of America*, 85(6), 2608-2622.
- Fox, C. M., Morrison, C. E., Ramig, L. O., & Sapir, S. (2002). Current perspectives on the lee silverman voice treatment (LSVT) for individuals with idiopathic parkinson disease. *American Journal of Speech Language Pathology*, 11(2), 111-123.
- Franke, R. H., & Kaul, J. D. (1978). The hawthorne experiments: First statistical interpretation. *American Sociological Review*, , 623-643.

- Gelb, D. J., Oliver, E., & Gilman, S. (1999). Diagnostic criteria for parkinson disease. *Archives of Neurology*, 56(1), 33-39.
- Gilkerson, J., & Richards, J. A. (2009). *The power of talk. impact of adult talk conversational turns and TV during the critical 0-4 years of child development.* (). Boulder, CO: LENA Foundation.
- Goberman, A. M., & Coelho, C. (2002). Acoustic analysis of parkinsonian speech II: L-dopa related flucturations and methodological issues. *NeuroRehabilitation*, 17(3), 247-254.
- Hammen, V. L., & Yorkston, K. M. (1996). Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria. *Journal of Communication Disorders*, 29(6), 429-445.
- Hammer, M. J., & Barlow, S. M. (2010). Laryngeal somatosensory deficits in parkinson's disease: Implications for speech respiratory and phonatory control. *Experimental Brain Research*, 201(3), 401-409.
- Harel, B., Cannizzaro, M., & Snyder, P. J. (2004). Variability in fundamental frequency during speech in prodromal and incipient parkinson's disease: A longitudinal case study. *Brain and Cognition*, 56(1), 24-29.
- Hartelius, L., & Svensson, P. (1994). Speech and swallowing symptoms associated with parkinson's disease and multiple sclerosis: A survey. *Folia Phoniatica Et Logopaedica*, 46(1), 9-17.
- Hellman, L. (2012). *Comparison of speech amplitude and intelligibility testing with measures of speech communication in daily life in parkinson disease.* (Unpublished Bachelor of Science). James Madison University,

- Herd, C. P., Tomlinson, C. L., Deane, K. H., Brady, M. C., Smith, C. H., Sackley, C. M., & Clarke, C. E. (2012a). Comparison of speech and language therapy techniques for speech problems in parkinson's disease. *The Cochrane Library*,
- Herd, C. P., Tomlinson, C. L., Deane, K. H., Brady, M. C., Smith, C. H., Sackley, C. M., & Clarke, C. E. (2012b). Speech and language therapy versus placebo or no intervention for speech problems in parkinson's disease. *The Cochrane Library*,
- Hillman, R. E., Heaton, J. T., Masaki, A., Zeitels, S. M., & Cheyne, H. A. (2006a). Ambulatory monitoring of disordered voices. *Annals of Otology, Rhinology & Laryngology*, 115(11), 795-801.
- Hillman, R. E., Montgomery, W. W., & Zeitels, S. M. (1997). Current diagnostics and office practice: Appropriate use of objective measures of vocal function in the multidisciplinary management of voice disorders. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 5(3), 172-175.
- Hillman, R. E., Heaton, J. T., Masaki, A., Zeitels, S. M., & Cheyne, H. A. (2006b). Ambulatory monitoring of disordered voices. *The Annals of Otology, Rhinology, and Laryngology*, 115(11), 795-801.
- Ho, A. K., Bradshaw, J. L., & Iannsek, R. (2008). For better or worse: The effect of levodopa on speech in parkinson's disease. *Movement Disorders*, 23(4), 574-580.
- Ho, A. K., Iannsek, R., & Bradshaw, J. L. (2002). The effect of a concurrent task on parkinsonian speech. *Journal of Clinical and Experimental Neuropsychology*, 24(1), 36-47.
- Ho, A. K., Bradshaw, J. L., & Iannsek, R. (2000). Volume perception in parkinsonian speech. *Movement Disorders*, 15(6), 1125-1131.

- Ho, A. K., Iannsek, R., & Bradshaw, J. L. (1999). Regulation of parkinsonia speech volume: The effect of interlocuter distance. *Journal of Neurology, Neurosurgery, and Psychiatry*, 67, 199-202.
- Ho, A. K., Bradshaw, J. L., Iannsek, R., & Alfredson, R. (1999). Speech volume regulation in parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia*, 37(13), 1453-1460.
- Huber, J. E., & Darling, M. (2011). Effect of parkinson's disease on the production of structured and unstructured speaking tasks: Respiratory physiologic and linguistic considerations. *Journal of Speech, Language, and Hearing Research*, 54(1), 33-46.
- Hunter, E. J. (2009). A comparison of a child's fundamental frequencies in structured elicited vocalizations versus unstructured natural vocalizations: A case study. *International Journal of Pediatric Otorhinolaryngology*, 73(4), 561-571.
- Hustad, K., & Weismer, G. (2007). Interventions to improve intelligibility and communicative success for speakers with dysarthria. *Motor Speech Disorders*, , 217-228.
- Jacobson, B. H., A., J., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI): Development and validation. *American Journal of Speech Language Pathology*, 6, 66-70.
- Kalf, J. G., Borm, G. F., de Swart, B. J., Bloem, B. R., Zwarts, M. J., & Munneke, M. (2011). Reproducibility and validity of patient-rated assessment of speech, swallowing, and saliva control in parkinson's disease. *Archives of Physical Medicine and Rehabilitation*, 92(7), 1152-1158.

- Keintz, C. K., Bunton, K., & Hoit, J. D. (2007). Influence of visual information on the intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology*, 16(3), 222-234.
- Kempler, D., & Lancker, D. V. (2002). Effect of speech task on intelligibility in dysarthria: A case study of parkinson's disease. *Brain and Language*, 80(3), 449-464.
- Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, 18(2), 124-132.
- Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders*, 54(4), 482-499.
- Kent, R., Kent, J., Weismer, G., Martin, R., Sufit, R., Brooks, B., & Rosenbek, J. (1989). Relationships between speech intelligibility and the slope of second-formant transitions in dysarthric subjects. *Clinical Linguistics & Phonetics*, 3(4), 347-358.
- Kim, Y., Kent, R. D., & Weismer, G. (2011). An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journal of Speech, Language, and Hearing Research*, 54(2), 417-429.
- Lee, C. (2001). Levodopa-induced dyskinesia: Mechanisms and management. *British Columbia Medical Journal*, 43(4), 206-209.
- 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, 47-57.

- Marsden, C., & Parkes, J. (1977). Success and problems of long-term levodopa therapy in parkinson's disease. *The Lancet*, 309(8007), 345-349.
- McAllister, A., & Brandt, S. K. (2012). A comparision of recordings of sentences and spontaneous speech: Perceptual and acoustic measures in preschool children's voices. *Journal of Voice*, 26(5), 673-e1.
- McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007a). The hawthorne effect: A randomised, controlled trial. *BMC Medical Research Methodology*, 7(1), 30.
- McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007b). The hawthorne effect: A rdomised, controlled trial. *BMC Medical Research Methodology*, 7(1)
- McHenry, M. (2011). An exploration of listener variability in intelligibility judgments. *American Journal of Speech-Language Pathology*, 20(2), 119-123.
- Mehta, D. D., Zanartu, M., Feng, S. W., Cheyne, H. A., & Hillman, R. E. (2012). Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *Biomedical Engineering, IEEE Transactions On*, 59(11), 3090-3096.
- Miller, N., Allcock, L., Jones, D., Noble, E., Hildreth, A. J., & Burn, D. J. (2007). Prevalence and pattern of perceived intelligibility changes in parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 78(11), 1188-1190. doi:jnnp.2006.110171 [pii]
- Morris, M. E., Iansek, R., Matyas, T. A., & Summers, J. J. (1994). The pathogenesis of gait hypokinesia in parkinson's disease. *Brain*, 117, 1169-1181.



- Muller, E. A. (1970). Influence of training and of inactivity on muscle strength. *Archives of Physical Medicine and Rehabilitation*, 51(8), 449-462.
- Oliveira, R. M., Gurd, J. M., Nixion, P., Marshall, J. C., & Passingham, R. E. (1997). Micrographia in parkinson's disease: The effect of providing external cues. *Journal of Neurology, Neurosurgery, and Psychiatry*, 63, 429-433.
- Oxtoby, M. (1982). *Parkinson's disease patients and their social needs: A survey of patients in contact with the parkinson's disease society* Parkinson's Disease Society.
- Pick, H. L., & al., e. (1989). Inhibiting the lombard effect. *The Journal of the Acoustical Society of America*, 85(2), 894-900.
- Pyo, H. Y., & Sim, H. S. (2007). A study of speech intelligibility affected by voice quality degradation. *Communication Sciences & Disorders*, 12(2), 256-278.
- Ramig, L. O. (1992). The role of phonation in speech intelligibility: A review and preliminary data from patients with parkinson's disease. *Intelligibility in Speech Disorders: Theory, Measurement and Management*, , 119-155.
- Ramig, L. O., Sapir, S., Fox, C., & Countryman, S. (2001). Changes in vocal loudness following intensive voice treatment (LSVT®) in individuals with parkinson's disease: A comparison with untreated patients and normal age-matched controls. *Movement Disorders*, 16(1), 79-83.
- Ramig, L. O., Countryman, S., Thompson, L. L., & Horii, Y. (1995). Comparison of two forms of intensive speech treatment for parkinson disease. *Journal of Speech, Language, and Hearing Research*, 38(6), 1232-1251.

- Ramig, L. O., & Dromey, C. (1996). Aerodynamic mechanisms underlying treatment-related changes in vocal intensity in patients with parkinson disease. *Journal of Speech, Language, and Hearing Research, 39*(4), 798-807.
- Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'Brien, C., Hoehn, M., & Thompson, L. L. (2001). Intensive voice treatment (LSVT) for patients with parkinson's disease: A 2 year follow up. *Journal of Neurology, Neurosurgery, and Psychiatry, 71*(4), 493-498.
- Rietveld, A. (2012). LJ beijer, MA, and RP clapham, MA. *Journal of Medical Speech-Language Pathology, 20*(2), 17-34.
- Rosen, K. M., & Duffy, J. R. (2006). Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers. *Journal of Speech, Language, and Hearing Research, 49*(2), 395-411.
- Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. *The Laryngoscope, 114*(9), 1549-1556.
- Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology, 22*(2), 212-226.
- Sadagopan, N., & Huber, J. E. (2007). Effects of loudness cues on respiration in individuals with parkinson's disease. *Movement Disorders, 22*, 651-659.
- Sáenz-Lechón, N., Godino-Llorente, J. I., Osma-Ruiz, V., Blanco-Velasco, M., & Cruz-Roldán, F. (2006). Automatic assessment of voice quality according to the GRBAS scale. Paper

presented at the *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, 2478-2481.

Sapir, S., Ramig, L. O., & Fox, C. M. (2011). Intensive voice treatment in parkinson's disease: Lee silverman voice treatment.

Sapir, S., Ramig, L., & Fox, C. (2008). Speech and swallowing disorders in parkinson disease. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 16(3), 205-210.  
doi:10.1097/MOO.0b013e3282febd3a [doi]

Sapir, S., Ramig, L. O., Hoyt, P., Countryman, S., O'Brien, C., & Hoehn, M. (2002). Speech loudness and quality 12 months after intensive voice treatment (LSVT) for parkinson's disease: A comparison with an alternative speech treatment. *Folia Phoniatrica Et Logopaedica : Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, 54(6), 296-303. doi:66148 [pii]

Sarno, M. T. (1968). Speech impairment in parkinson's disease. *Archives of Physical Medicine and Rehabilitation*, 49(5), 269-275.

Schloneger, M., & Hunter, E. J. (2015). Correlations between simultaneous voice qualify readings by accelerometer and acoustic transducers. *11th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research*, London, United Kingdom.

Schulz, G. M., & Grant, M. K. (2000). Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in parkinson's disease: A review of the literature. *Journal of Communication Disorders*, 33(1), 59-88.

- Schulz, G. (2004). The effects of pallidotomy surgery on sentence measures across three tasks in parkinson patients. *Journal of Medical Speech-Language Pathology*, 12(4), 195-205.
- Sewall, G. K., Jiang, J., & Ford, C. N. (2006). Clinical evaluation of parkinson's-Related dysphonia. *The Laryngoscope*, 116(10), 1740-1744.
- Smith, M. E., Ramig, L. O., Dromey, C., Perez, K. S., & Samandari, R. (1995). Intensive voice treatment in parkinson disease: Laryngostroboscopic findings. *Journal of Voice*, 9(4), 453-459.
- Speyer, R., Bogaardt, H. C., Passos, V. L., Roodenburg, N. P., Zumach, A., Heijnen, M. A., . . . Brunings, J. W. (2010). Maximum phonation time: Variability and reliability. *Journal of Voice*, 24(3), 281-284.
- SPSS, I. (2012). IBM SPSS statistics version 21. *Boston, Mass: International Business Machines Corp*,
- Stathopoulos, E. T., Huber, J. E., Richardson, K., Kamphaus, J., DeCicco, D., Darling, M., . . . Sussman, J. E. (2014). Increased vocal intensity due to the lombard effect in speakers with parkinson's disease: Simultaneous laryngeal and respiratory strategies. *Journal of Communication Disorders*, 48, 1-17.
- Stelmach, G. E., Teasdale, N., Phillips, J., & Worringham, C. J. (1989). Force production characteristics in parkinson's disease. *Experimental Brain Research*, 76, 165-172.
- Taaffe, D. R., Duret, C., Wheeler, S., & Marcus, R. (1999). Once-weekly resistance exercise improves muscle strength and neuromuscular performance in older adults. *Journal of the American Geriatrics Society*, 47(10), 1208-1214.

- 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*,
- Tjaden, K. K., & Liss, J. M. (1995). The role of listener familiarity in the perception of dysarthric speech. *Clinical Linguistics & Phonetics*, 9(2), 139-154.
- van den Boogaart, K., Rost, S., Petzoldt, T., & Petzoldt, M. T. (2014). Package 'qualV'.
- Van Lancker Sidsis, D., Cameron, K., & Sidsis, J. J. (2012). Dramatic effects of speech task on motor and linguistic planning in severely dysfluent parkinsonian speech. *Clinical Linguistics & Phonetics*, 26(8), 695-711.
- Walsh, B., & Smith, A. O. (2012). Basic parameters of articulatory movements and acoustics in individuals with parkinson's disease. *Movement Disorders*, 27(7), 843-885.
- Weismer, G. (1984). Articulatory characteristics of parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination. *The Dysarthrias: Physiology, Acoustics, Perception, Management*, , 101-130.
- Weismer, G., Yunusova, Y., & Bunton, K. (2012). Measures to evaluate the effects of DBS on speech production. *Journal of Neurolinguistics*, 25(2), 74-94.
- Wheeler, K. M., Collins, S. P., & Sapienza, C. M. (2006). The relationship between VHI scores and specific acoustic measures of mildly disordered voice production. *Journal of Voice*, 20(2), 308-317.
- Wickström, G., & Bendix, T. (2000). The "hawthorne effect"—what did the original hawthorne studies actually show? *Scandinavian Journal of Work, Environment & Health*, , 363-367.

- Wight, S., & Miller, N. (2014). Lee silverman voice treatment for people with parkinson's: Audit of outcomes in a routine clinic. *International Journal of Language & Communication Disorders*,
- World Health Organization. (2002). *Towards a common language for functioning, disability and health: ICF*. World Health Organisation.
- Yorkston, K. M. (2010). *Management of motor speech disorders in children and adults* Pro-ed.
- Yorkston, K. M., & Beukelman, D. R. (1984). *Assessment of intelligibility of dysarthric speech* Pro-ed.
- Yorkston, K. M., Hakel, M., Beukelman, D. R., & 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.
- Yorkston, K. M., Beukelman, D. R., & Tice, R. (1996). Sentence intelligibility test. *Tice Technology Services*,
- Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America*, 71(6), 1544-1550.
- Zraick, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z., & Glaze, L. E. (2011). Establishing validity of the consensus auditory-perceptual evaluation of voice (CAPE-V). *American Journal of Speech-Langauge Pathology*, 20(1), 14-22.

**Appendix I:** List of potential conversational topics

- Home/Distance
  - How was your trip over?
  - Did it take long to get here?
  - Did you find it okay?
- Family
  - Do you have children?
  - Do they live nearby? Do you get to see them often?
- Parkinson's Disease
  - Tell me a little about how you found out about your Parkinson's Disease.
- Work
  - What do/did you do for work? Did you enjoy it?
- Hobbies
  - What do you like to do for fun?

**Manuscript: Effects of Environment and Task on Speech Functioning in Parkinson's Disease and Healthy Controls**

**Effects of Environment and Task on Speech Functioning in Parkinson's Disease and Healthy Controls**

**Carrie E. Rountrey, Nina M. Borrás, Christy L. Ludlow** James Madison University  
Department of Communication Sciences and Disorders, Harrisonburg, Virginia, USA

Address for Correspondence: Carrie E. Rountrey (and other authors),  
Communication Sciences and Disorders, James Madison University, 801 Carrier  
Drive, MSC 4304, Harrisonburg, Virginia USA 22807. Emails: [rountrce@jmu.edu](mailto:rountrce@jmu.edu),  
[nmborras@dukes.jmu.edu](mailto:nmborras@dukes.jmu.edu), [ludlowcx@jmu.edu](mailto:ludlowcx@jmu.edu)



**Abstract**

**Objective:** To compare the effects of clinical and home environments on speech in persons with Parkinson's disease (PD) and healthy controls (HV).

**Participants:** Thirteen persons with PD and speech difficulty and a control group of twelve healthy volunteers with normal speech were recorded.

**Methods:** A small portable digital recorder was used to record spontaneous conversations of participants in both settings. Spontaneous speech intelligibility and intensity were measured from sentences recorded during interactions in a clinic setting and selected randomly from 15 hours of recordings made in the home.

**Results:** The PD group had reduced intelligibility compared to HV in both environments ( $p=0.01$ ) with no differences found between the clinic and home environments. Groups did not differ in speech intensity but both had reduced intensity in the clinic compared with home environment ( $p=0.02$ ). This difference was not correlated with environmental noise or articulation rate. Spontaneous speech intelligibility in the clinic was related to spontaneous speech intelligibility in the home environment ( $R^2=0.309$ ) in the PD group.

**Conclusions:** PD patients' speech intelligibility was equal to that measured in the clinic, was decreased compared to healthy volunteers, and may reflect disease. Possibly motivation to communicate with significant others contributed to speech intensity increases in the home and may not reflect disease.

## Introduction

Symptoms of Parkinson's disease (PD) include resting tremor, reduced range of movement (rigidity) and slow movement (bradykinesia) all of which affect motor functions such as walking, speech and writing(1). All of the speech subsystems (respiratory, laryngeal, and supralaryngeal) can be affected in persons with PD(2-7). Persons with PD usually suffer from hypokinetic dysarthria, which includes reduced voice loudness, monoloudness, monopitch, disordered rate, disordered articulation, hoarse, breathy, or harsh voice and/or tremulous voice (8,9). Symptoms may change to hyperkinesia 1-hour post medication in some patients(10-12). Hypophonia accompanying PD is characterized by a soft voice and reduced intelligibility(8,13). Overall, 89% of those with PD have voice problems and 45% have articulation problems(9).

Reduced speech intensity has long been studied and treated in PD, as hypophonia is a hallmark of the disease in many people with PD(14-18). The most commonly utilized therapy focuses on vocal loudness for changes in intensity and posits that other speech subsystems benefit from the single task of increased loudness(19).

Intelligibility is defined as "the degree to which the speaker's intended message is recovered by the listener" (20) and may depend upon both communication task and environmental cues. Large differences in percent intelligibility between spontaneous speech and reading aloud have been reported(21). Significant differences between the intelligibility of single words, sentences, and monologues have also been reported in persons with PD and

dysarthria(22). Within the clinical environment (speech therapy office, rehab setting, laboratory), the emphasis on speech may help the patient to focus on their speech. For example, reading tasks may allow patients to concentrate solely on *how* they are speaking. Patients are aware when their speech is being recorded, whether at home or in the clinical environment, however the focus on speech in the clinical environment while reading sentences may differ from communicating in the home environment. Listeners have difficulty understanding those with dysarthric speech(23). In a controlled clinical environment when the patient is focused on performing to their maximum capability, their speech intelligibility may improve(24). Patients' perception of their speech intelligibility may differ from their listeners. Some have suggested that persons with PD may report that they have less severe intelligibility problems than their communication partners report from listening to them in the home environment(25-27). Improved speech intelligibility is the primary aim of treatment; however, this has not been measured as often as speech intensity after treatment for PD(28).

Differences in speech produced in the clinic versus speech communication in the natural or home environment for speakers with PD have been suggested in the literature(29-31). These differences may be due to differences in cognitive load, performance effects, competition for cognitive resources, and environmental cues in the two environments. In the home environment the speaker needs to focus on functional communication, which may increase cognitive load(32) and distract attention from speech. In 2006, Burton et al. (33) observed that speakers with PD were judged to be more intelligible during clinical tasks than in the natural

environment by their spouses. During a dual task when persons were asked to read sentences while screwing in a bolt with their non-dominant hand, the patients had lower sentence intelligibility scores, comparable to their spontaneous speech, compared to a speech-only condition. Environmental cues have been shown to improve motor and speech functions in PD(6,34,35).

Given the described differences between environments and that judgments are made clinically based on the clinical environment alone, exploration into the relationship between clinical measures and the home environment may have explanatory value and important implications for clinical practice.

Standard clinical tasks for evaluation of hypophonia and dysarthria in PD include both acoustic measures, such as: mean sound pressure level (SPL) on a prolonged vowel(19), range of SPL (SPL- $\Delta$ ) (36), maximum phonation time (MPT) (37), harmonic to noise ratio (HNR) (38), and range of  $F_0$ (39); and perceptual measures such as: the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) (40) and the Sentence Intelligibility Test (SIT) (41). Clinicians often include quality of life scales such as the Voice Handicap Index (VHI) (42,43).

SPL is usually measured by taking the mean intensity over the vocalized portion of the sustained vowel, usually /a/ or /i/(44), while  $F_0$ -range can be measured either in sentences or from pitch glides. As males and females differ in total  $F_0$  range, using the interquartile range over a target sentence is thought to correct for gender differences(45). MPT is measured from the vocal onset to offset during a sustained phonation task(46). HNR is calculated from sustained phonation and can be expressed in dB, for example if 99% of the energy in the signal is in

harmonics, and 1% is noise between harmonics, the HNR is  $10 \cdot \log_{10}(99/1) - 20\text{dB}$ (47).

Acoustic measurements such as SPL have been associated with increased intelligibility in patients with PD(44,48). Although HNR has not yet been directly studied in its relationship to intelligibility (PubMed search: “intelligibility AND HNR” reveals 5 articles, none with direct relationships investigated), the GRBAS scale, a perceptual measure of voice(49,50) has been correlated with intelligibility. The more severe the GRBAS rating, the lower the intelligibility, despite intact articulation and prosody(51). Perceptual measures such as the CAPE-V and the SIT are also common evaluative measures for persons with PD. The CAPE-V(40) is similar to GRBAS but was developed to standardize perceptual ratings of voice. As patients are rated on spontaneous conversation and several speech tasks, intelligibility may contribute to perceptual ratings. The Sentence Intelligibility Test (SIT) records patients reading sentences of increasing length and then generates a percent intelligibility based on three listeners’ transcriptions of those sentences(41). As the SIT uses connected speech tasks, it is thought to be relevant to spontaneous connected speech.

Task effects have been demonstrated within the clinical environment (5,52) (5,59) by changing situational cues and noting intensity changes(52), and between extemporaneous and read speech tasks, and by noting linguistic and respiratory changes(5). Clinical tasks have also been related to clinical ratings of speech clarity(53). Because information on how clinical tasks relate to intelligibility in the home were not found in the literature, the present study aimed to examine the

relationship between the most common clinical tasks and speech intelligibility in the home environment.

Intelligibility in the clinical environment is usually judged either informally during the interview process (spontaneous speech), using a perceptual rating scale like the CAPE-V (spontaneous speech given prompt), or formally using a measure like the SIT (reading task). CAPE-V overall score interrater reliability has been shown to be .76(54) and the SIT can show 15% variability between listeners(55). Still, these connected speech tasks and perceptual judgments are the closest representation that we have of what happens in the home environment between speaker and listener. No quantitative measures have reported on speech intelligibility in the home.

In this study, we sought to compare spontaneous speech during communication in the home environment with spontaneous speech in a clinical environment. It was hypothesized that both speech intensity and speech intelligibility would be greater in the clinical environment when compared to the home environment, in both persons with PD and healthy controls. It was further hypothesized that the greater difference between environments would be found in persons with PD. We also aimed to explore typical clinical measures and how they may relate to percent sentence intelligibility in the home environment ( $SI_{HOME}$ ).

## **Methods**

James Madison University's Institutional Review Board (IRB) and Sentara Rockingham Memorial Hospital's IRB approved the protocol for this study, including the consent and all recruitment materials.

### *Participants*

Participants were recruited through word-of-mouth, flyers, local PD support groups, and Sentara Rockingham Memorial Hospital Voice and Swallowing Services. Healthy controls were recruited in the same way and through email blasts to staff at James Madison University. A phone screening was used to determine inclusion in the study. Informed consent was sent to all included participants ahead of time for informational purposes, and was reviewed and signed at the time of their appointment.

Inclusion criteria for the PD group were: over 50 years of age, a score of 23 or higher was required on the Folstein Mini Mental State Examination (MMSE) (56), diagnosis of PD by a neurologist, speech/voice impairment typical of hypophonia and/or dysarthria commensurate with PD as judged by a practicing speech-language pathologist with experience in neurological disorders, and passing an audiological screening at 40 dB-HL at 500, 1000, 2000, 3000 and 4000 Hz (allowing for a 4000 Hz notch in one ear, if the participant had experience with firearms). Exclusion criteria for the PD group included: history of voice or speech problems prior to PD diagnosis, diagnosis of neurological disease(s) other than PD, current diagnosis of mental illness except depression, and a non-native Standard American English speaker.

Inclusion criteria for the HV group were identical with the exception of having the PD diagnosis. Exclusion criteria for the HV group were identical except that they were without an adult history of speech problems.

Participants with PD were encouraged to make their appointments at a time of day that was best for them, given their experience with their medication schedule. HV made appointments at their convenience. Both groups tended to choose mid- to late- morning appointments, usually within 1-2 hours after medication. No participants with PD exhibited dyskinesias at the time of the testing.

### *Listeners*

Listeners were recruited from communication sciences and disorders classes at James Madison University. Listeners volunteered their time and gave informed consent. Inclusion criteria for the listener group included: Normal hearing (able to pass audiological screening at 30 dB-HL for 500, 1K, 2K, 3K, and 4K Hz), and native speaker of Standard American English. Exclusion criteria were: presence of disease, delay or disorder affecting listening or attending to task.

### *Speech Recording*

Clinical recordings were made in a sound attenuated room. Average background noise was 35 dB-SPL, measured using a Quest Model 2200 sound level meter in the center of the room. The entire session was recorded using the LENA digital language processor (DLP) (57) placed 9 inches from the mouth (Fig. 1). The DLP contains an omnidirectional microphone with a flat 20-20 kHz frequency response. Frequencies above 10 kHz are suppressed, as they are unlikely to contain human speech. Low frequencies are suppressed by a 70 Hz high-pass filter. Digital data were filtered using a 10 kHz low-pass filter to suppress high-frequency sounds. Acoustic waveforms were recorded using a 16 kHz 16-bit sigma-delta analog to digital (ADC) converter with 8x over-sampling digital interpolation.



The LENA DLP has a hardware limiter for intensity suppression at 87dbSPL. This limiter was not problematic for this study, given that normal conversational speech averaged 79 dB-SPL with the mic-to-mouth distance set at 9 inches. No clipping was observed in the sentence samples that were gathered. When tested on pure tones at various frequencies, the DLP showed an equal flat response at 125 Hz, 250 Hz, 500 Hz, 750 Hz, 1000 Hz, 1500 Hz, 2000 Hz, 3000 Hz, 4000 Hz, 6000 Hz, and 8000 Hz.

The DLP was calibrated using 500 Hz calibration tones at 65, 70, 75, 80 and 85 dB-HL. The tones were delivered via Grason-Stadler GSI 61 clinical audiometer and Grason-Stadler speakers in a sound attenuated room. The DLP and a sound level meter (set at 50 dB low, C-weight, and fast speed) were both 9 inches from the speaker. The SPL was read on the sound level meter while calibration recordings were saved for linear interpolation to convert into dB-SPL using Praat(58).

The slopes and y-intercepts derived from the linear interpolation for each of the two DLPs used for the study were: (DLP 1, slope=1.0447, y-intercept=9.146,  $R^2=.9997$ ; DLP 2, slope=1.0447, y-intercept=12.784,  $R^2=.9997$ ).

The DLP was clipped on participants' chests 9 inches below their mouth, to record spontaneous speech and the assessment tasks in the clinical setting (Fig. 1).

### *Speech Tasks*

Clinical sessions began with reviewing consent, placing the LENA-DLP on the participant at 9 inches from their mouth to start recording, followed by casual conversation. Topics were either initiated by the participant or taken from a list of topics that the researcher initiated, such as, "Did you have to travel far today?" and,

“Does your family live nearby? Do you get to see them often?” This recording of spontaneous conversation provided a clinical recording of at least 5 minutes that could be contrasted with speech conversation in the home.

The Voice Handicap Index (VHI) (42,43) and Beck Depression Inventory (BDI) (59) were administered as paper and pencil tests. Participants then participated in a hearing screening and the MMSE (56). Participants were excluded if they did not pass the hearing screening at 40 dB-SPL (allowing for high frequency loss at 40K Hz in one ear), and/or if they didn't score at least 23 points on the MMSE. Instructions were given for formal speech testing, which included the Sentence Intelligibility Test (41) and a script recording of speech tasks. The script recording included tasks and examples to produce a sustained vowel and MPT. To measure  $F_0$  range in sentences, two phrases were imitated, “Say, that’s excellent,” and “No, he meant you.” The script included other tasks not measured for the present study.

After speech testing was completed, the participants were instructed on the LENA-DLP, recording instructions. Participants were given written instructions for device use and a letter to inform any other conversational partners that they may encounter during the study about the recording. They were instructed on how to switch the device on and off should they or their conversational partner wish to have a private conversation. They were instructed not to use the DLP in the presence of others not aware that the device was recording or did not agree to have their speech recorded.

All of the clinical recording sessions were collected by the same experimenter at a time of day that the patient determined was “best” for them, given their experience with their own medication schedule.

### *Recordings in the Home Environment*

Recordings in the natural environment were made in the participants’ home and other non-clinical settings in which their primary conversational partner was usually their spouse, with prior agreement of the partner. Participants were instructed on maintaining a 9-inches microphone-to-mouth distance and given a ribbon to maintain the same mouth-to-mic distance over the two days of recording. They were also provided with written instructions on making recordings with the DLP. Participants then completed 15 hours of recording within their home environment over two days.

Spontaneous utterances from the clinical session and the home environment were reduced to conversational segments, time stamped, and selected via a random number generator. Spontaneous utterances were selected from the sound file by matching the random number generator to timestamps on the recording. The nearest sentence within a conversation that met the following criteria was selected: From intentional 1-on-1, person-to-person communication, between 4 and 15 words long without background noise interfering with speech and a unique sentence. Sentences were excluded if: the conversation was with more than one person; the conversational partner was not aware they were being recorded; the sentence was too long for listeners’ memory during transcription (>15 words); or was a repeated sentence. Ten randomly selected sentences from each environment

were saved in a digital audio file (.wav) for further intelligibility and intensity analysis.

Audio files of clinical tasks included: sustained vowel at comfortable pitch and loudness, sustained vowel for maximum phonation time, sentences “Say, that’s excellent,” and “No, he meant you,” SIT sentences, and a compilation of sound samples to be used for CAPE-V ratings.

#### *Data Analysis: Intelligibility*

Three different novel, inexperienced listeners transcribed each speaker’s spontaneous sentences and sentences from the SIT, with no one listener hearing the same sentence more than once. Each was limited to 20 minutes of listening and transcription to reduce listener fatigue. Listeners used laptops to transcribe what they heard as the sentences were played for them. Prior to playing each set of sentences, amplification levels on the speakers were set to match the same sound level meter values while playing back the calibration tones recorded during the clinical session. Listeners were instructed that they would hear each sentence only once and that they were to transcribe exactly what they heard, without regard for punctuation or capitalization. They were told not to make any corrections until the entire sentence was transcribed, to reduce their memory load. Pauses were given until the listener was ready to start the next sentence and after each set of 10 sentences.

Sentences were transcribed into Microsoft Excel spreadsheets and compared with transcriptions made by the researchers, who had listened to the sentences in context for an unlimited number of times and, when needed, at a higher intensity.

Comparisons were made with R, Package 'qualV' (60) using the Least Common Sequence (LCS) function algorithm. Prior to running the program, the entire spreadsheet was checked for spelling errors. The LCS algorithm was adapted to do a number of things: text was converted to lowercase without punctuation, single spaces between words and delete leading spaces, so as not to count these as transcriber errors. The algorithm then created output columns that examined for the least common sequence, or the number of identical words that were in the same sequence as the key sentence. Then, the number of identical words in the sequence was divided by the greatest number of words in either sentence, so that errors of omission and commission were accounted for. The mean percent intelligibility of a set of 10 sentences for each spontaneous task, and 11 for the SIT task was determined and saved in the dataset as  $SI_{HOME}$  (home percent speech intelligibility),  $SI_{CLINIC}$  (clinic percent speech intelligibility) and  $SIT_R$  (SIT as calculated by R).

To determine the accuracy of the LCS algorithm, a random sample of 10 spontaneous sentences from  $SI_{HOME}$  and  $SI_{CLINIC}$  were calculated for intelligibility by hand based on an early SIT protocol(61), then compared to the LCS code written for R, resulting in a perfect correlation ( $R^2=1.0$ ) with 0% disagreement. Comparisons were also made between the R LCS output for SIT sentences ( $SIT_R$ ) and the SIT software output(41) for percent word intelligibility in sentences in a sample of 44 transcribed SIT sentences, resulting in a very high correlation ( $R^2=.98$ ) with 3.26% disagreement. R LCS allowed for over 3500 lines of ineligibility to be analyzed in moments, reducing bias and human error.

*Data Analysis: Acoustic Measures*

All acoustic measurements (SPL, HNR, MPT,  $F_0$  mean and standard deviation) were analyzed using Praat (47) software for speech analysis version 5.3.70.

Intensity of each spontaneous utterance was measured with Praat intensity set to show a range of 25dB to 100dB with “subtract mean pressure” checked as OFF. If pauses in the midst of an utterance were greater than 150 ms, they were deleted (62) before the mean intensity was computed, from the beginning to the end of the sound contour (Fig. 2). The intensity level was converted into decibels in sound pressure level re .0002 dynes/cm<sup>2</sup> using the slope and y-intercept of the linear interpolation derived from measurement of calibration tones from the LENA-DLP used to record the sample. The mean dB-SPL of each participant’s sentences in each environment was calculated ( $SPL_{CLINIC}$  and  $SPL_{HOME}$ ) by measuring the intensity of each randomly selected utterance and calculating the average dB-SPL of the 10 utterances from the clinic and the average dB-SPL of the 10 utterances from home.

Intensity of a sustained vowel was calculated in the same fashion. SPL measure was taken from the middle 90% of the sustained vowel /a/ at comfortable loudness and pitch, and the same settings and procedures as above ( $SPL_{VOWEL}$ ).

Harmonic to Noise Ratio (HNR) was also taken from this sustained vowel at comfortable loudness and pitch.

Maximum Phonation Time (MPT) was measured from participants response the script task “now I want you to say /a/ and hold it for as long as you can.” The voiced signal was visualized, the cursor was set at the beginning and the end of the sound signal, and the time was noted and saved for the dataset.

Frequency range was measured from participants' imitations of the auditory models presented via digital recording for, "Say, that's excellent," and "No, he meant you." Praat was used to detect the mean fundamental frequency and standard deviation from each sentence. and the interquartile range was calculated via the 75<sup>th</sup> percentile z-score (0.674) to find the actual 75<sup>th</sup> percentile of the distribution. Z-score  $0.674 = (x_{75} - \text{mean}) / (\text{SD})$ , and solve for  $x_{75}$  for the 75<sup>th</sup> percentile. Then, because of the symmetry of the standard normal distribution, the 25<sup>th</sup> percentile is calculated as the same distance from the mean as the 75<sup>th</sup> percentile, but in the opposite direction;  $x_{25} = \text{mean} - (x_{75} - \text{mean})$ . The difference between the 25<sup>th</sup> and 75<sup>th</sup> ( $x_{75} - x_{25} = \text{IQR}$ ) percentile was saved for the dataset ( $F_0 - \text{IQR}$ ).

#### *Data Analysis: Perceptual Measures*

CAPE-V ratings were performed by an expert voice clinician, who was given sound samples for each participant ( $n=13$ ), including distractor samples of healthy controls ( $n=12$ ). Each file was randomly numbered so that the identification of each participant was masked and not identified as either PD or HV. Each digital recording included: 1-2 minutes of conversation between the experimenter and participant, SIT sentences, and the script recording of speech tasks. The expert clinician listened to the digital recordings and filled out a CAPE-V rating form on each. The experimenter then measured the overall rating and calculated the overall score for each participant (out of 100 mm).

#### *Statistical Methods*

A power analysis was conducted based on Hellman, 2012(63). Effect size differences between the clinic and home measures ranged from 0.38 to 1.47 for

intelligibility and from 0.41 to 1.8 for intensity in four participants. For a power of 0.80, 11 subjects were required to find differences on measures of speech intelligibility at  $\alpha=0.025$  while 8 subjects were required to find intensity differences at  $\alpha=0.025$  (Bonferroni corrected for 2 outcome variables) between clinic and home environments.

Two-way repeated measures ANOVA compared clinic versus natural environment (environment) within subjects and PD versus healthy volunteers (group) between subjects and the interaction of environment and group for each dependent variable: speech intelligibility, and intensity.

Exploratory stepwise multiple regression analyses examined the predictor variables  $SIT_R$ , CAPE-V,  $SPL_{VOWEL}$ , HNR, MPT and  $F_0$ -IQR against the outcome variable  $SI_{HOME}$ . Single regressions were also performed on each predictor variable.

## Results

### *Participants*

Thirteen participants with idiopathic Parkinson's disease (7 Men, 6 women, 55 to 74 years, median 70 years) participated in the study. The 12 healthy volunteers were over 50 (4 Men, 8 women, 53 to 70 years, median 60.5 years). The participant data are presented in **Table 1**. All but one participant was on medication for PD, the one who chose holistic treatments; all earned a minimum score of 23/30 on the MMSE(56). The range of overall impairment at the time of speech evaluation extended from mild to moderate severe disability. Three participants had deep brain stimulators and settings remained constant during the study. Of the 13 participants with PD, 7 reported no history of speech therapy, 4



reported Lee Silverman Speech Therapy (LSVT), or “4 weeks of talking loud,” and 2 reported other types of speech therapy. The most recent speech therapy was 1.5 years prior to participation in the study. Two participants in the HV group were spouses of volunteers in the PD group. Out of 27 volunteers, two did not participate: One volunteer withdrew prior to participation as their primary communication partner in the home opted not to participate; another did not have any speech or voice difficulties.

### *Listeners*

Fifty-eight listeners participated in the study. Their ages ranged from 18 to 53 years (mean 21.9 years). None had completed clinical experience or coursework in motor speech disorders or voice at the time of their participation. All listeners were native speakers of standard American English, free of neurological or psychiatric disorders, had visual acuity adequate to perform the task, and passed a hearing screening. Sixty volunteered, one was excluded due to not passing the hearing screening, and another was excluded after self-reporting an auditory processing disorder. Most listeners did one session per day ( $n=53$ ), and no more than three sessions per day ( $n=3$ ).

### *Inter-rater reliability*

All listeners' responses on 3 of the same samples were examined using intraclass correlation coefficients. Intraclass correlations were 0.946-0.980, exceeding our predetermined minimum standard of 0.65.

### *Effect of Environment*

Statistical analyses were completed using SPSS professional statistics version 21(64). Two-way repeated ANOVAs compared environment (clinical and home) within subjects and between groups (PD vs. controls).

Significant effects of environment were found on intensity ( $F=12.563, p=.02$ ). No group difference was found ( $F=1.274, p=0.271$ ) and no interaction between intensity and group ( $F=.025, p=.86$ ). Specifically, both groups had higher mean intensity levels in the home environment than in the clinical environment (Fig. 3). Environmental effects within groups were significant; HV ( $F=7.624, p=0.019$ ) and PD ( $F=5.694, p=0.034$ ).

Intelligibility scores did not differ between environments ( $F=.389, p=.539$ ) with no group by environment interaction ( $F=.048, p=.829$ ). However, significant group differences were found across both environments ( $F=14.988, p=0.001$ ), and in the clinic ( $F=13.552, p=0.001$ ) and the home environment ( $F=10.055, p=0.004$ ) (Fig. 4).

#### *Correlation Coefficients Between Clinical Measures and Speech in the Home*

Correlation coefficients examined relationships between speech intensity and intelligibility in the clinic and the home environment. No significant relationships ( $p \leq 0.05$ ) were found between intensity and intelligibility within either group in either environment. Within the PD group, a significant relationship was found between spontaneous speech intelligibility in the clinical and the home environments ( $R^2=.309, p=.048$ ). No significant relationships were found between years since diagnosis and intensity in the home environment ( $R^2=.297, p=.054$ ); between years since diagnosis and intelligibility in the home environment ( $R^2=.259$ ,

$p=.076$ ); between speech intensity in the home environment and the depression index ( $R^2=.125$ ,  $p=.236$ ) or between intelligibility in the home environment and depression ( $R^2=.118$ ,  $p=.250$ ). The average background noise within 50 ms of target sentences in the home environment and average sentence dB-SPL shared a weak, non-significant, positive correlation ( $R^2=0.289$ ,  $p=0.058$ ) across participants.

### *Exploring Predictions Between Clinical Measures and Speech in the Home*

To explore the relationship between percent speech intelligibility in the home and measures of speech in the clinic in a person with PD (Table 2), we performed a stepwise multiple regression. The combination with the highest predictive value was  $SPL_{VOWEL}$  and  $SIT_R$  in relationship with  $SI_{HOME}$   $F(2,10)=12.755$ ,  $p=.002$ . The multiple correlations coefficient for  $SIT_R+SPL_{VOWEL}$   $R^2=.662$ , indicating that approximately 66.2% of the variance of  $SI_{HOME}$  could be accounted for by  $SIT_R+SPL_{VOWEL}$ . From this data, the prediction equation can be constructed as:

$$\text{Predicted } SI_{HOME} = 83.901 + 1.683(SIT_R) - 1.922(SPL_{VOWEL})$$

This indicated that the  $SIT_R$  score was adjusted downward by  $SPL_{VOWEL}$  to best predict  $SI_{HOME}$ .

Single regressions of each predictor variable against the outcome variable,  $SI_{HOME}$ , while correcting alpha for 6 variables ( $p<.008333$ ), revealed that only  $SIT_R$   $F(1,11)=10.546$ ,  $p=.008$  had a significant prediction of  $SI_{HOME}$  ( $R^2=0.489$ ,  $p=.004$ ) (Fig.5). No other variable had predictability alone.

## **Discussion**

This study compared intensity and intelligibility between the clinical and home environments in PD and healthy volunteers. Our expectation was that within

the clinic, the contextual emphasis on speech might focus the participants more on *how* they are speaking, while at home, participants would be more distracted and focus on *what* they are going to say and not *how* they are speaking, thereby reducing intelligibility and intensity. We also expected patients would be distracted from communication by concurrent tasks (32,65) in the home that were not present in the clinic. By using the same device and conversational speech in both settings, we sought to minimize testing effects.

Statistical analysis of intelligibility and intensity within the clinical and home environments showed different effects from those expected. Both groups had higher speech intensity in the home environment. Individuals' speech intensity did not correlate with their intelligibility scores and no environmental effects were found in speech intelligibility. The results of the current study show differences between environments only in speech intensity. As this did not relate to background noise, the increased intensity in the home environment is likely due to other factors, such as motivation and proximity to communication partner.

Although it is common clinical practice is to address intensity in speech therapy for PD and hypophonia is an agreed-upon marker of the disease (66), we found that speech intensity did not differ between healthy volunteers and participants with PD. Persons with PD had lower intelligibility scores than the HV in both environments but did not differ in intensity from HV in either environment. Previous research has focused on clinical outcomes of therapy for voice/speech problems in PD, focusing on intensity measures such as dB-SPL on a sustained vowel (19). This study demonstrates that an increased focus on intelligibility as a

measure of speech impairment in PD is warranted. However, it should be noted that this was an exploratory study based on natural speech intensity and intelligibility. The participants were not given any instruction to increase their loudness or effort during the time of this investigation. As there was no manipulation of speech intensity in this study, no conclusions can be made whether intensity enhancement might alter speech intelligibility in the home environment.

Previous authors had suggested that the communication partners of PD patients report speech intelligibility at home was more problematic than in the clinic (25-27). Historically, clinicians and researchers have observed differences in the clinic versus the natural environments (29,30,33) suggesting that cognitive load and testing effects may lead to differences between “performance” speech and “functional” speech. Our data did not support these clinical impressions. Performance effects have also been discussed as possible enhancers of speech in the clinical environment (67), that is a speech-language pathologist may encourage participants to have clearer and/or louder speech. However, results of this study indicate that people with PD had greater speech intensity in their home environment but no difference in intelligibility. This indicates that intensity of spontaneous speech is affected by environment, in patients with PD but did not differ from HV. Spontaneous speech intelligibility was not affected by environment, but was affected by PD in comparison with HV. This suggests that intelligibility of spontaneous speech may more reflective of the effect of PD on speech communication (26). As the intensity of spontaneous speech did not differ between PD and HV, it is less likely to be a measure of PD effects on speech. When

considering measures of speech impairment in PD, spontaneous speech intelligibility should hold a primary position in clinical measures, as it reflected speech intelligibility in the home. New methods of effective and efficient measurement of intelligibility that can be used in both environments are needed.

To explore the relationship between the environments in the PD sample, we examined predictability of clinical measures used in speech and voice clinical evaluations with the speech intelligibility of the same participants in their home environments. Two variables that best fit a prediction model for  $SI_{HOME}$  were combined  $SIT_R$  and  $SPL_{VOWEL}$ . These findings suggest that clinical speech tasks that do not require connected speech (such as sustained vowels for intensity, harmonic to noise ratio, and maximum phonation time) were not predictive of intelligibility in the home. This is contrary to previous research indicating that intensity, maximum phonation time and frequency range are associated with clinical intelligibility (44,48). However, other findings are in agreement with more current research indicating that intensity interquartile range was not significantly predictive of clinical intelligibility scores (53).

Our results showed that sentence intelligibility in a reading task ( $SIT_R$ ) corrected by intensity of a sustained vowel ( $SPL_{VOWEL}$ ) best predicted  $SI_{HOME}$ . Perhaps task must be considered when attempting to draw connections between clinical tasks and functional speech tasks in the clinic. The  $SIT_R$  score derived from a reading task was not related to spontaneous speech intelligibility in the clinic. Further,  $SPL_{VOWEL}$  was not related to intelligibility in either environment. Perhaps voice and speech tasks that do not call on the patient with PD to use connected

speech, such as isolated vowel tasks, may not be good indicators of speech impairment in PD. It has been indicated that learning and behavioral effects may blur these tasks over time(68).

Although the model of  $SIT_R + SPL_{VOWEL}$  was statistically significant, accounting for 66.2% of the variance in  $SI_{HOME}$ , 33.8% remains unaccounted for and with the significant individual predictor, the  $SIT_R$ , there was still 51.1% of the variance in  $SI_{HOME}$  left unaccounted for. Thus, clinical tests used here could not substitute for measuring speech intelligibility in the home environment. Further research is warranted on clinical protocols and tasks that may relate to spontaneous speech in patients' functional environments. Kim, Kent, and Weismer (53) reported that when acoustic measures are regressed against scaled intelligibility scores (mild, moderate, severe) on a reading task, in several disease groups (PD, Stroke, Traumatic Brain Injury, Multiple Systems Atrophy) that some measures had predictive power to intelligibility, including articulation rate ( $R^2=.56$ ), and  $F_2$  slope ( $R^2=.51$  for males,  $R^2=.46$  for females). The PD group in the Kim et. al study had the higher intelligibility ratings, and therefore may have different results as a group. Currently, reading intelligibility is used to rate intelligibility in the clinic. As discussed, the  $SIT_R$  score does not seem to be closely related to  $SI_{CLINIC}$  ( $R^2=.21$ ), but  $SIT_R$  was related to  $SI_{HOME}$  ( $R^2=.49$ ).

Our results suggest that clinical measures of voice and speech may only partially indicate the speech communication abilities of patients with PD in the home. The  $SIT$  and measures of  $SPL$  in combination can lend some insight into the speech behavior of the patient at home. The  $SIT_R$  was the only measure to indicate

spontaneous speech intelligibility in the home. This exploratory study warrants further exploration in a larger, more diverse (with regard to impairment), group of patients.

Limitations of this study include a relatively small sample size, and that the majority of participants with PD were only mildly to moderately affected. Also, the use of clinical measures in common practice may not be as tightly controlled as our measures were in this study. For example, our measure  $SIT_R$ , was well controlled with regard to calibration, listener experience, and processing the data through R LCS. This alternative to the commercially available version gave us a higher level of control. We were able to demonstrate a very high level of correlation to the commercially available SIT computer scoring program ( $R^2=.98$ ,  $err.=3.26\%$ ) although some score differences may occur.

Further exploration is needed into contributions of background noise, motivation, distance from conversational partners, and other cues from conversational partners. For our study, we examined the relationship between the mean environmental noise of all subjects in home and clinical environments, with the mean speech intensities in each environment and did not find a significant relationship. Comparing within-subjects possible changes in sentence intensity in response to changes in immediate background noise may show different results. Motivation to be understood by functional communication partners in the home may contribute positively to spontaneous speech intensity in the home environment. In our experiment, conversational proximity was held constant in the clinical environment, but could not be held constant in the home environment, and could



have contributed to greater intensity at times in the home. Also, other cues from conversational partners, such as asking for clarification, off topic responses, or even lack of response, were not measured in this study.

This study emphasized the importance of the home environment in which patients have most of their communication and where their communication is important to quality of life. Clinical tasks cannot approximate the differing cognitive loads and competition for resources in the home environment. New technologies are making this type of measurement increasingly possible. Hardware is getting smaller and more discreet, and software is becoming more powerful and portable. In this study we utilized the LENA-DLP, a simple encrypted digital recorder. Other hardware options that are available are voice dosimeters, accelerometers (69) and the combinations of accelerometers and acoustic transducers (70). Software for voice analysis can be loaded on smartphones (71), and speech transcription software may be applied to intelligibility. In the future, ambulatory monitoring of speech and voice is expected to become more efficient and effective.

Clinically, this study suggests that spontaneous speech intelligibility is a major factor affected in PD, and not spontaneous speech intensity. We found that participants who were louder, even in spontaneous speech, did not necessarily have greater speech intelligibility. This study examined several factors that may be related to spontaneous speech intelligibility for communication at home. Even among the strongest relationship of  $SIT_R$  ( $R^2=.49$ ) to  $SI_{HOME}$ , this did not explain more than half of the variance. Therefore, future research and development should

include efficient and effective methods of measurement of intelligibility in the home environment.

### **Clinical Messages**

- Spontaneous speech intelligibility deficits in the clinical and home environments are related.
- Perceived environmental effects from communication partners and professionals must be corroborated by quantitative measures.
- Spontaneous speech intensity in mild to moderately severe PD may not differ from healthy volunteers, while intelligibility is impaired in PD regardless of environment.

### **Contributors**

CER participated in the original initiation and design of the study, devised procedures, took part in recruitment, data collection, processing and analysis. She was the principal investigator and the lead author in preparing drafts.

NMB participated in data collection and processing, lab organization, and commenting on draft versions of the work.

CLL participated in the original initiation and design, data analysis, and reviewing and critically commenting on drafts and the final version.

### **Acknowledgements**

The authors gratefully acknowledge funding from James Madison University department of Communication Sciences and Disorders, which supported CER as a Ph.D. candidate and NMB as a master's student during this time. Purdue University

also paid mileage to CER for patients who enrolled in the treatment study running in tandem with this study.

**Competing Interests**

We have no competing interests to report.

## References

- (1) Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. *Arch Neurol* 1999;56(1):33-39.
- (2) Connor NP, Abbs JH, Cole KJ, Gracco VL. Parkinsonian deficits in serial multiarticulate movements for speech. *Brain* 1989;112(4):997-1009.
- (3) Forrest K, Weismer G, Turner GS. Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults. *J Acoust Soc Am* 1989;85(6):2608-2622.
- (4) Duffy JR. Motor speech disorders: Substrates, differential diagnosis, and management. : Elsevier Health Sciences; 2012.
- (5) Huber JE, Darling M. Effect of Parkinson's disease on the production of structured and unstructured speaking tasks: Respiratory physiologic and linguistic considerations. *Journal of Speech, Language, and Hearing Research* 2011;54(1):33-46.
- (6) Sadagopan N, Huber JE. Effects of loudness cues on respiration in individuals with Parkinson's disease. *Movement Disorders* 2007;22:651-659.
- (7) Hammer MJ, Barlow SM. Laryngeal somatosensory deficits in Parkinson's disease: implications for speech respiratory and phonatory control. *Experimental brain research* 2010;201(3):401-409.
- (8) Darley FL, Aronson AE, Brown JR. Clusters of deviant speech dimensions in the dysarthria. *J Speech Hear Res* 1969b;12:462-496.
- (9) Logemann JA, Fisher HB, Boshes B, Blonsky ER. Frequency and Cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *J Speech Hear Disord* 1978;43:47-57.
- (10) Lee C. Levodopa-induced dyskinesia: Mechanisms and management. *British Columbia Medical Journal* 2001;43(4):206-209.
- (11) Schulz GM, Grant MK. Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in Parkinson's disease: a review of the literature. *J Commun Disord* 2000;33(1):59-88.
- (12) Marsden C, Parkes J. Success and problems of long-term levodopa therapy in Parkinson's disease. *The Lancet* 1977;309(8007):345-349.
- (13) Darley FL, Aronson AE, Brown JR. Differential diagnostic patterns of dysarthria. *J Speech Hear Res* 1969a;12:246-269.
- (14) Ramig LO, Sapir S, Fox C, Countryman S. Changes in vocal loudness following intensive voice treatment (LSVT®) in individuals with Parkinson's disease: A comparison with untreated patients and normal age-matched controls. *Movement Disorders* 2001;16(1):79-83.

- (15) Sapir S, Ramig L, Fox C. Speech and swallowing disorders in Parkinson disease. *Curr Opin Otolaryngol Head Neck Surg* 2008 Jun;16(3):205-210.
- (16) Sapir S, Ramig LO, Hoyt P, Countryman S, O'Brien C, Hoehn M. Speech loudness and quality 12 months after intensive voice treatment (LSVT) for Parkinson's disease: a comparison with an alternative speech treatment. *Folia Phoniatr Logop* 2002 Nov-Dec;54(6):296-303.
- (17) Wight S, Miller N. Lee Silverman Voice Treatment for people with Parkinson's: audit of outcomes in a routine clinic. *International Journal of Language & Communication Disorders* 2014.
- (18) Herd CP, Tomlinson CL, Deane KH, Brady MC, Smith CH, Sackley CM, et al. Comparison of speech and language therapy techniques for speech problems in Parkinson's disease. *The Cochrane Library* 2012.
- (19) Ramig LO, Sapir S, Countryman S, Pawlas AA, O'Brien C, Hoehn M, et al. Intensive voice treatment (LSVT) for patients with Parkinson's disease: a 2 year follow up. *J Neurol Neurosurg Psychiatry* 2001 Oct;71(4):493-498.
- (20) Kent R, Kent J, Weismer G, Martin R, Sufit R, Brooks B, et al. Relationships between speech intelligibility and the slope of second-formant transitions in dysarthric subjects. *Clinical Linguistics & Phonetics* 1989;3(4):347-358.
- (21) Kempler D, Lancker DV. Effect of speech task on intelligibility in dysarthria: a case study of Parkinson's Disease. *Brain Lang* 2002;80(3):449-464.
- (22) Bunton K, Keintz CK. The use of a Dual-Task Paradigm for Assessing Speech Intelligibility in Clients with Parkinson Disease. *J Med Speech Lang Pathol* 2008 Sep 1;16(3):141-155.
- (23) Tjaden KK, Liss JM. The role of listener familiarity in the perception of dysarthric speech. *Clinical Linguistics & Phonetics* 1995;9(2):139-154.
- (24) McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne Effect: a randomised, controlled trial. *BMC medical research methodology* 2007;7(1).
- (25) Kalf JG, Borm GF, de Swart BJ, Bloem BR, Zwarts MJ, Munneke M. Reproducibility and validity of patient-rated assessment of speech, swallowing, and saliva control in Parkinson's disease. *Arch Phys Med Rehabil* 2011;92(7):1152-1158.
- (26) Miller N, Allcock L, Jones D, Noble E, Hildreth AJ, Burn DJ. Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2007 Nov;78(11):1188-1190.
- (27) Ho AK, Bradshaw JL, Iansek R. Volume perception in Parkinsonian speech. *Movement Disorders* 2000;15(6):1125-1131.
- (28) Deane KH, Ellis-Hill C, Jones D, Whurr R, Ben-Shlomo Y, Playford ED, et al. Systematic review of paramedical therapies for Parkinson's disease. *Movement Disorders* 2002;17(5):984-991.

- (29) Sarno MT. Speech impairment in Parkinson's disease. *Arch Phys Med Rehabil* 1968 May;49(5):269-275.
- (30) Weismer G. Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination. *The dysarthrias: Physiology, acoustics, perception, management* 1984:101-130.
- (31) Keintz CK, Bunton K, Hoit JD. Influence of visual information on the intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology* 2007;16(3):222-234.
- (32) Dromey C, Benson A. Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research* 2003;46(5):1234-1246.
- (33) Effects of a concurrent motor task on speech intelligibility for speakers with Parkinson disease. *Conference on Motor Speech*, Austin, TX; 2006.
- (34) Oliveira RM, Gurd JM, Nixion P, Marshall JC, Passingham RE. Micrographia in Parkinson's disease: the effect of providing external cues. *J Neurol Neurosurg Psychiatr* 1997;63:429-433.
- (35) Ho AK, Bradshaw JL, Iansek R, Alfredson R. Speech volume regulation in Parkinson's disease: effects of implicit cues and explicit instructions. *Neuropsychologia* 1999;37(13):1453-1460.
- (36) Buder EH, Cannito MP. Acoustic assessment of vocal function. *Neurologic disorders of the larynx*: Thieme Medical Publishers New York; 2009. p. 74-84.
- (37) Sewall GK, Jiang J, Ford CN. Clinical Evaluation of Parkinson's-Related Dysphonia. *Laryngoscope* 2006;116(10):1740-1744.
- (38) Predicting severity of Parkinson's disease from speech. *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE: IEEE*; 2010.
- (39) Harel B, Cannizzaro M, Snyder PJ. Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: a longitudinal case study. *Brain Cogn* 2004;56(1):24-29.
- (40) Kempster GB, Gerratt BR, Abbott KV, Barkmeier-Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *American Journal of Speech-Language Pathology* 2009;18(2):124-132.
- (41) Yorkston KM, Beukelman DR, Tice R. Sentence intelligibility test. *Tice Technology Services* 1996.
- (42) Jacobson BH, A. J, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The voice handicap index (VHI): development and validation. *American Journal of Speech Language Pathology* 1997;6:66-70.
- (43) Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. *Laryngoscope* 2004;114(9):1549-1556.

- (44) Dromey C, Ramig LO, Johnson AB. Phonatory and Articulatory Changes Associated With Increased Vocal Intensity in Parkinson Disease: A Case Study. *Journal of Speech, Language, and Hearing Research* 1995;38(4):751-764.
- (45) Baken RJ, Orlikoff RF. *Clinical measurement of speech and voice*. : Cengage Learning; 2000.
- (46) Speyer R, Bogaardt HC, Passos VL, Roodenburg NP, Zumach A, Heijnen MA, et al. Maximum phonation time: variability and reliability. *Journal of Voice* 2010;24(3):281-284.
- (47) Boersma P, Weenink D. {P} raat: doing phonetics by computer. 2010.
- (48) Ramig LO, Countryman S, Thompson LL, Horii Y. Comparison of two forms of intensive speech treatment for Parkinson disease. *Journal of Speech, Language, and Hearing Research* 1995;38(6):1232-1251.
- (49) De Bodt MS, Wuyts FL, Van de Heyning, Paul H, Croux C. Test-retest study of the GRBAS scale: influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice* 1997;11(1):74-80.
- (50) Automatic assessment of voice quality according to the GRBAS scale. *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE: IEEE; 2006*.
- (51) Pyo HY, Sim HS. A Study of Speech Intelligibility Affected by Voice Quality Degradation. *Communication Sciences & Disorders* 2007;12(2):256-278.
- (52) Ho AK, Iannsek R, Bradshaw JL. Regulation of parkinsonia speech volume: the effect of interlocuter distance. *J Neurol Neurosurg Psychiatr* 1999;67:199-202.
- (53) Kim Y, Kent RD, Weismer G. An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journal of Speech, Language, and Hearing Research* 2011;54(2):417-429.
- (54) Zraick RI, Kempster GB, Connor NP, Thibeault S, Klaben BK, Bursac Z, et al. Establishing validity of the consensus auditory-perceptual evaluation of voice (CAPE-V). *American Journal of Speech-Language Pathology* 2011;20(1):14-22.
- (55) McHenry M. An exploration of listener variability in intelligibility judgments. *American Journal of Speech-Language Pathology* 2011;20(2):119-123.
- (56) Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12(3):189-198.
- (57) Ford M, Baer CT, Xu D, Yapanel U, Gray S. The LENATM language environment analysis system: Audio specifications of the DLP-0121 2008.
- (58) Boersman P, Weenink D. Praat: doing phonetics by computer [Computer program].

- (59) Beck AT, Steer RA, Brown GK. Beck depression inventory-II. San Antonio 1996.
- (60) van den Boogaart K, Rost S, Petzoldt T, Petzoldt MT. Package 'qualV'. 2014.
- (61) Yorkston KM, Beukelman DR. Assessment of intelligibility of dysarthric speech. : Pro-ed; 1984.
- (62) Hammen VL, Yorkston KM. Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria. *J Commun Disord* 1996;29(6):429-445.
- (63) Hellman L. Comparison of Speech Amplitude and Intelligibility Testing with Measures of Speech Communication in Daily Life in Parkinson Disease. 2012.
- (64) SPSS I. IBM SPSS statistics version 21. Boston, Mass: International Business Machines Corp 2012.
- (65) Ho AK, Iansek R, Bradshaw JL. The effect of a concurrent task on Parkinsonian speech. *Journal of Clinical and Experimental Neuropsychology* 2002;24(1):36-47.
- (66) Herd CP, Tomlinson CL, Deane KH, Brady MC, Smith CH, Sackley CM, et al. Speech and language therapy versus placebo or no intervention for speech problems in Parkinson's disease. *The Cochrane Library* 2012.
- (67) Hustad K, Weismer G. Interventions to improve intelligibility and communicative success for speakers with dysarthria. *Motor speech disorders* 2007:217-228.
- (68) Rietveld A. LJ Beijer, MA, and RP Clapham, MA. *Journal of Medical Speech-Language Pathology* 2012;20(2):17-34.
- (69) Hillman RE, Heaton JT, Masaki A, Zeitels SM, Cheyne HA. Ambulatory monitoring of disordered voices. *Annals of Otolaryngology, Rhinology & Laryngology* 2006;115(11):795-801.
- (70) Correlations Between Simultaneous Voice Quality Readings by Accelerometer and Acoustic Transducers. 11th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research; 2015.
- (71) Mehta DD, Zanartu M, Feng SW, Cheyne HA, Hillman RE. Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *Biomedical Engineering, IEEE Transactions on* 2012;59(11):3090-3096.



***Table 1: Participant Clinical Data***

ID	Age	Gender	Voice Handicap Index	Beck Depression Index	Years since Diagnosis	MMSE
<b><i>Clinical Data of Participants with Parkinson's Disease</i></b>						
301	72	F	50	19	10	27
302	74	M	40	3	10	28
303	71	M	67	16.5	4	29
304	55	F	59	9	2.5	30
305	71	M	42	3	6	28
306	70	M	37	0	13	29
307	63	F	52	0	10	27
308	65	M	60	17	15	30
309	57	M	8	9	2	29
310	70	M	65	2	3.5	27
311	73	M	18	2	2	29
312	55	F	50	10	5	29
313	74	M	48	30	6	30
<b><i>Clinical Data of Healthy Volunteers</i></b>						
401	56	F	3	.	.	30
402	59	F	9	14	.	30
403	53	M	2	1	.	30
404	55	F	11	.	.	29
405	62	M	3	6	.	30
406	57	F	7	0	.	29
407	70	F	4	3	.	25
408	63	F	0	0	.	30
409	69	M	9	0	.	30
410	65	M	16	2	.	29
411	63	F	0	4	.	29
412	56	F	1	0	.	28

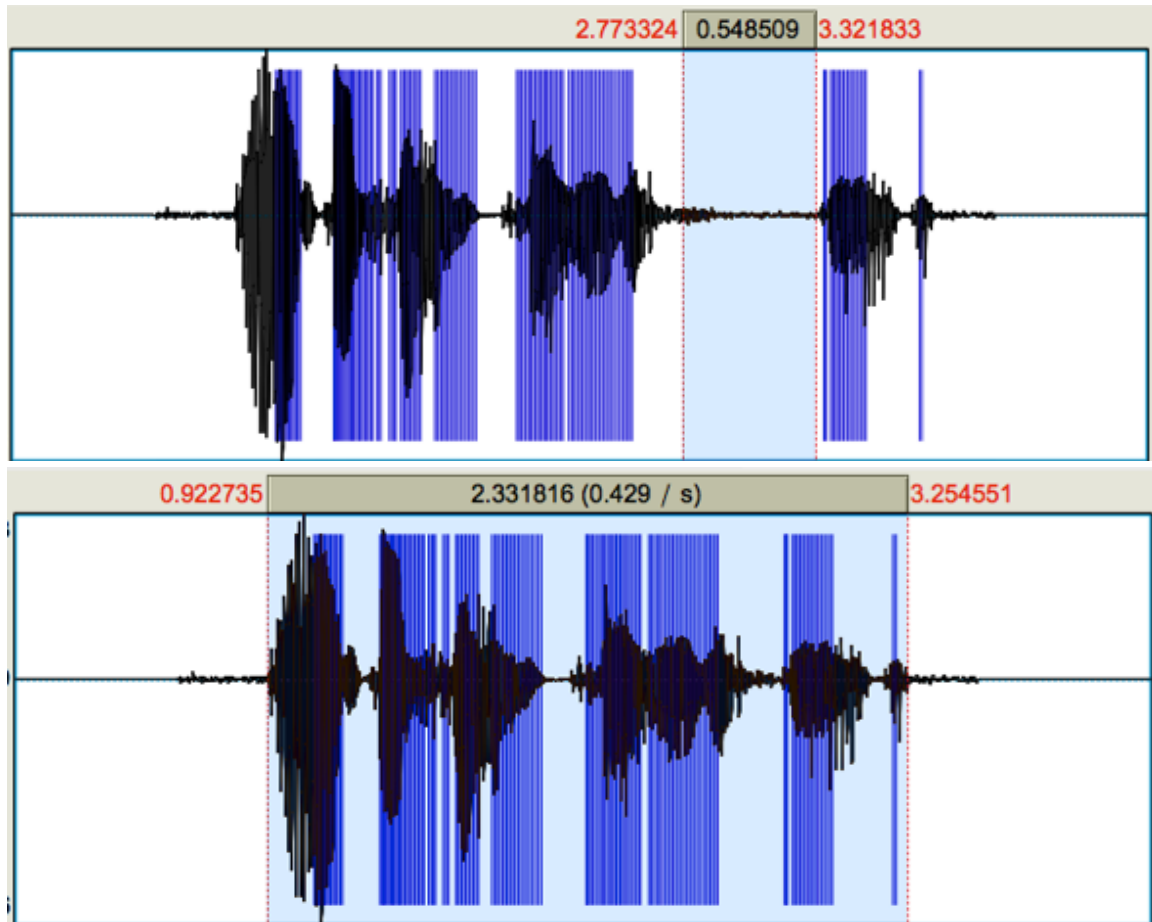
***Table 2: Examination of Means (n=13) for patients with PD***

	<i>Mean</i>	<i>Standard Deviation</i>
Home Percent Sentence Intelligibility ( $SI_{HOME}$ )	72.46	17.30
Consensus Auditory Perceptual Evaluation of Voice (CAPE-V)	25.23	12.51
Sentence Intelligibility Test ( $SIT_R$ )	86.23	9.88
Max Phonation Time (MPT)	10.76	3.41
Harmonic to Noise Ratio (HNR)	18.53	4.72
Fundamental Frequency Interquartile Range ( $F_0$ -IQR)	43.63	19.39
Sound Pressure Level in dB-SPL ( $SPL_{VOWEL}$ )	81.43	4.91

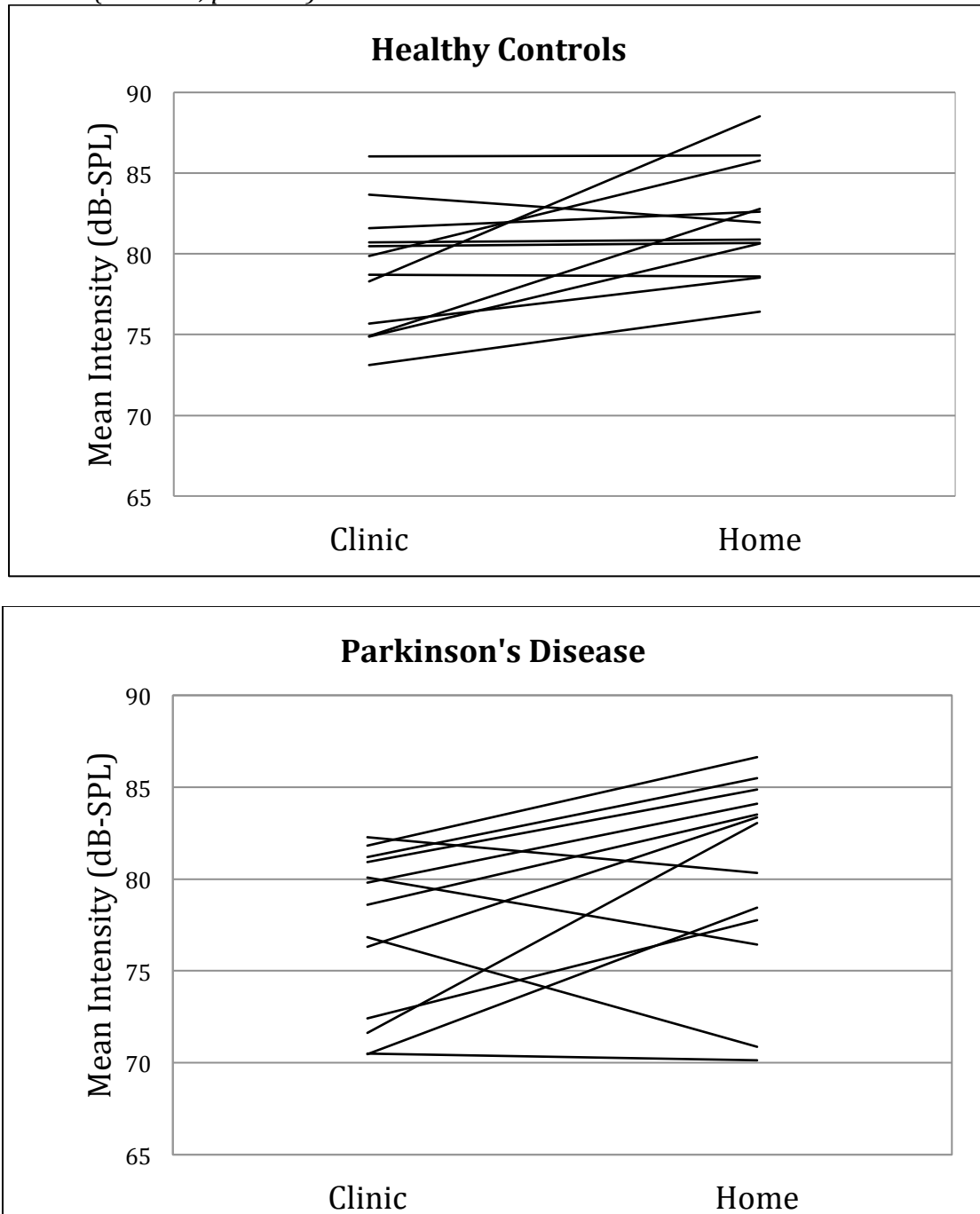
**Figure 1: LENA-DLP.** The DLP is 2.5oz, measures approximately 8cm x 5cm x 1cm. Shown here is the money clip holder, used to hold the DLP and worn on model that is measuring the distance between mouth and microphone on DLP, as participants were instructed.



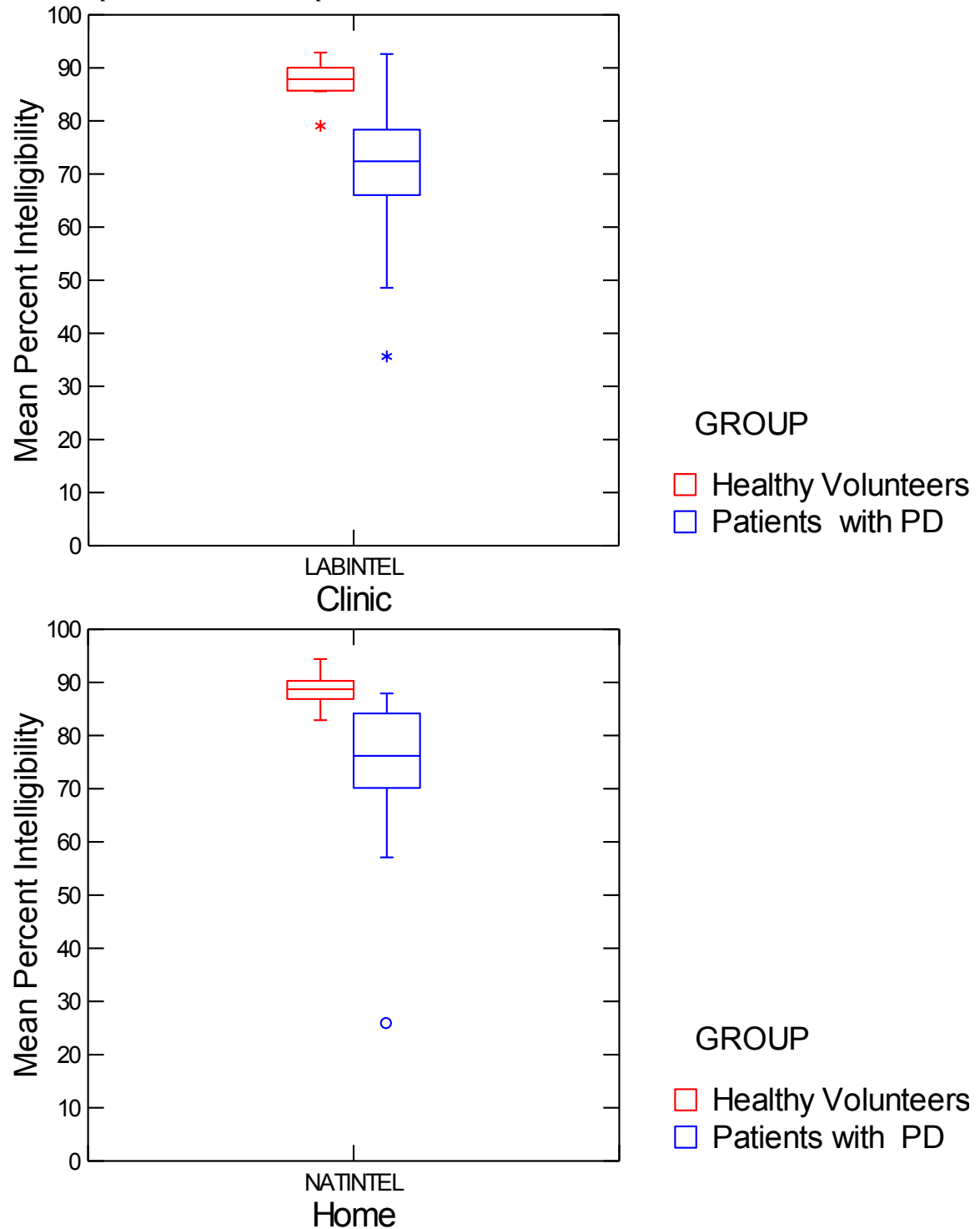
**Figure 2: Data Processing for Intensity.** Demonstration of processing the utterance, “More or less keeping my brain a little more active,” intensity reading in Praat, with pause (selected 0.5485 seconds in top image) and with pause removed (bottom). Bottom image shows selection of utterance (length indicated as 2.33 seconds) and deselection of noise before/after target utterance. Once the utterance is selected, mean intensity appears alongside the spectrograph, or can be selected through dropdown menus.



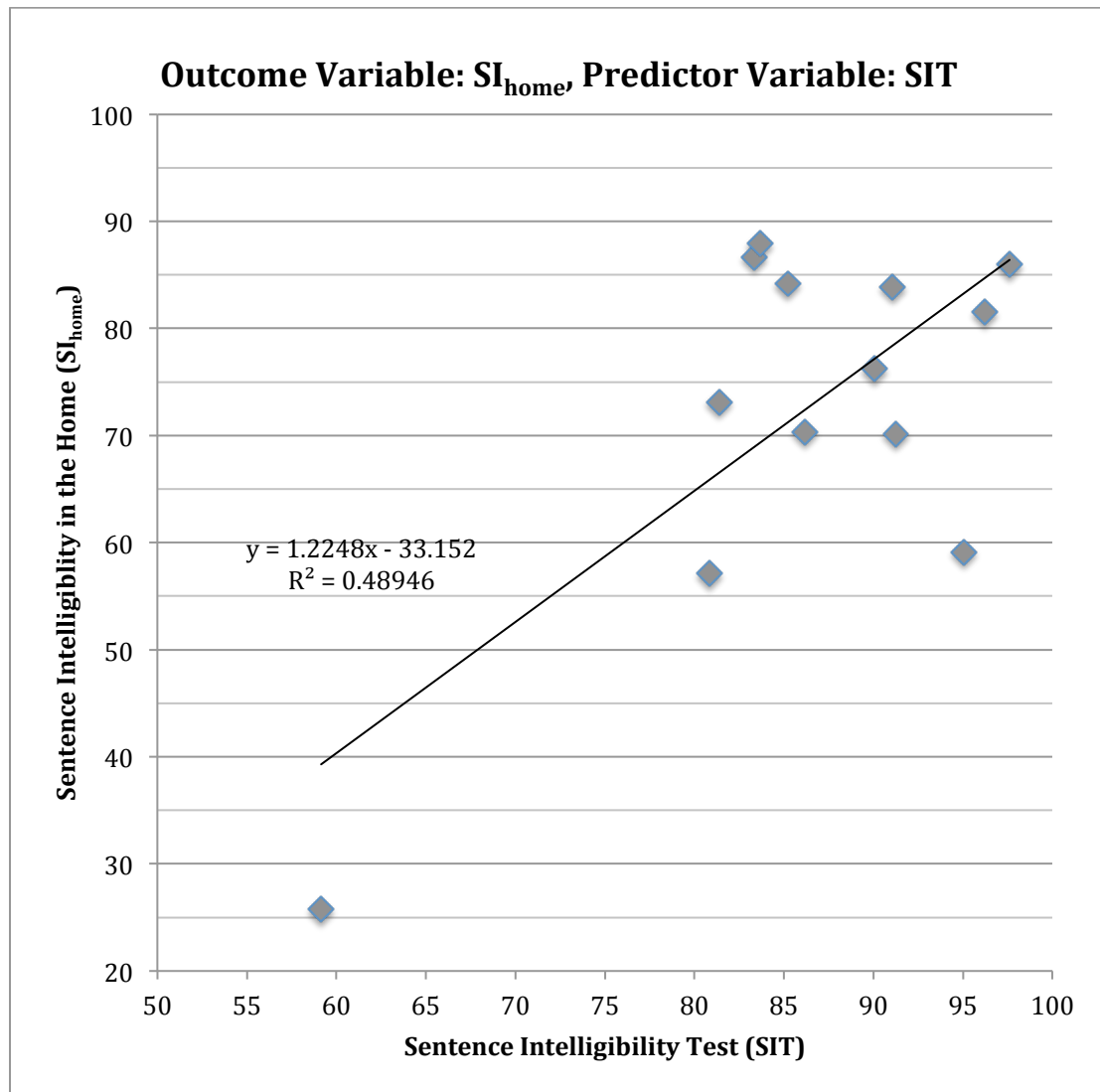
**Figure 3: Examination of Intensity Between Groups.** Showing intensity within environments clinic and home for healthy controls (top) and PD (bottom). Significant effects of environment ( $F=12.563$ ,  $p=.02$ ). No group difference ( $F=1.274$ ,  $p=0.271$ ) and no interaction between intensity and group ( $F=.025$ ,  $p=.86$ ). Specifically, both groups had higher mean levels of intensity in the home environment than in the clinical environment. Environmental effects were found within groups, HV ( $F=7.624$ ,  $p=0.019$ ) and PD ( $F=5.694$ ,  $p=0.034$ ).



**Figure 4: Spontaneous Speech Percent Intelligibility by Environment and Group.**  
Boxes represent 1<sup>st</sup> and 2<sup>nd</sup> quartiles and median.



***Figure 5: Regression of Sentence Intelligibility Test on Percent Sentence Intelligibility in the Home***



**Pilot Study: Application of Ambulatory Recording Device for Measurement of a 12-week Treatment Protocol**

**The Effect of Use of SpeechVive™ on Unaided Intelligibility and Intensity in the Home Environment in Parkinson's Disease: A Pilot Study**

Carrie E. Rountrey, Nina M. Borrás, Jessica E. Huber, Christy L. Ludlow

James Madison University

Department of Communication Sciences and Disorders



## **Abstract**

**Purpose:** To assess pre- and post- treatment measures of speech intelligibility and intensity, when unaided, within the natural environment, before and after 12 weeks of treatment with the SpeechVive™ wearable device.

**Methods:** Seven patients with PD were recorded in their homes with the LENA-DLP wearable device when not using the SpeechVive™ device. Spontaneous sentences were randomly chosen for analysis. Three different listeners judged sentence intelligibility. Intensity of spontaneous sentences was interpolated via calibration of devices with a sound pressure meter.

**Results:** Repeated measures ANOVAs determined that before and after application of SpeechVive™ device for 12 weeks, there was no significant difference in intelligibility ( $F(1,6)=.071, p=.798$ ) or intensity ( $F(1,6)=.021, p=.890$ ). The effect sizes of Cohen's  $d$  were 0.055 and 0.15 respectively. Given the very small effect sizes, these pilot data suggest that the effects on unaided speech are negligible after using the device for 12 weeks.

**Conclusions:** Application of the SpeechVive™ device for 12 weeks did not elicit a statistically significant difference in spontaneous speech intelligibility or intensity in the home, when not wearing the aid. Further investigation with a larger sample size is not warranted. However, the sample included only mild to moderately affected patients who may be less susceptible to treatment effect.

## Introduction

In recent years, many persons with PD who seek treatment for hypophonia and reduced intelligibility do so through Lee Silverman Voice Treatment (LSVT), but not all patients have demonstrated progress using LSVT. Generalization of skills to the natural environment has not been demonstrated, and this may be due to sensory and cognitive problems in persons with PD (Countryman & Ramig, 1993; Fox et al., 2002). LSVT requires 4 days/week for 4 weeks of intensive therapy and compliance with homework to be successful, and also requires the person with PD to remember to “think loud,” when s/he is in functional communication situations. Many of the participants in LSVT’s published clinical trials are persons with PD who have mild to moderate dysarthria and relatively intact cognition (Ramig et al., 1995; Ramig & Dromey, 1996).

The effects of external cues on limb movement and voice in persons with PD may indicate that people with PD might benefit from targets for motor movements, both gross (limb) and fine (speech/voice). Such cues have resulted in improvement in handwriting, gait and voice loudness (Ho et al., 1999; Ho et al., 1999; Morris et al., 1994; Oliveira et al., 1997; Sadagopan & Huber, 2007).

To alleviate possible difficulties with self cuing apparent in PD, and to provide a naturalistic external cue, Huber and colleagues have developed the SpeechVive™, a wearable device for persons with PD. The device fits on the ear like a behind-the-ear hearing aid (Fig.1) and delivers “babble,” to elicit the Lombard effect when the patient speaks. The device senses the patient’s vocalization through an accelerometer and only delivers background “babble” when the patient is

speaking. Sadagopan and Huber (2007) have shown similar responses to multi-talker babble in persons with PD and healthy controls. Both groups increased their sound pressure level and respiratory kinematics resulting in increases in intensity and efficiency of respiratory patterns with multi-talker babble. As the SpeechVive™ device can be used in the home, it can become a part of the patient's activities of daily living, which should improve generalization of skills and heighten the probability of realizing gains in communication at home.

Principles of practice and exercise physiology indicate that heavier loading on the muscles results in increased strength (Taaffe et al., 1999). As patients speak more loudly, they may be increasing strength in the respiratory muscles thus the SpeechVive™ may increase speech intensity after 12 weeks of speaking with the device. If speech intensity increases with device use, then speech intelligibility may increase as well (Forrest et al., 1989). Therefore we hypothesized that intelligibility in connected speech would increase after 12 weeks of device use. We measured this using the LENA digital encrypted recording device (Ford et al., 2008) in the home environments of the participants, without the device on to determine if there is carryover in the patients' functional environment.

## **Methods**

### **Participants**

All of the participants also participated in Rountrey et al., in preparation. Participants were recruited through word-of-mouth, flyers, local PD support groups, and Sentara Rockingham Memorial Hospital Voice and Swallowing Services. A phone screening was used to determine inclusion in the study. Informed consent

was sent to all included participants ahead of time for informational purposes, and was reviewed and signed at the time of their appointment. James Madison University's Institutional Review Board (IRB) and Sentara Rockingham Memorial Hospital's IRB approved the protocol for this study, including informed consent and all recruitment materials.

Our inclusion criteria dictated that our participants with PD must have a diagnosis from a neurologist of PD, and must be 50 years of age or older. They were to earn a score of 23 or higher on the Folstein Mini Mental State Examination (MMSE) (Folstein et al., 1975), and pass an audiological screening at 40 dB-HL at 500, 1000, 2000, 3000 and 4000 Hz (allowing for 4000 Hz notch in one ear, especially if the participant had experience with firearms). The participants' speech/voice impairment had to be typical of hypophonia and/or dysarthria commensurate with PD as judged by a practicing speech-language pathologist with experience in neurological disorders. Exclusion criteria for the study included: history of voice or speech problems prior to PD diagnosis, diagnosis of neurological disease(s) other than PD, current diagnosis of mental illness except depression, and non-native Standard American English speaker.

Participants were encouraged to make their initial appointments at a time of day that was best for them, given their experience with their medication schedule. Participants tended to choose mid- to late- morning appointments, usually within 1-2 hours after medication. No participants exhibited dyskinesias at the time of initial testing, although dyskinesias during the home environment recordings can not be ruled out.

## **Listeners**

Listeners were recruited from classes in Communication Sciences and Disorders at James Madison University. Listeners volunteered their time and gave informed consent. Inclusion criteria for the listener group included normal hearing for age (able to pass audiological screening at 30 dB-HL for 500, 1K, 2K, 3K, and 4K Hz), and native speaker of Standard American English. Exclusion criteria included presence of disease, delay or disorder affecting listening or attending to task.

## **Speech Recording**

Speech recordings were made in the home environment of each participant and the pre-session data is the same as was reported in Rountrey et al., in preparation. Both pre- and post- sessions were recorded using the LENA digital language processor (DLP) (Ford et al., 2008) placed 9 inches from the mouth (Fig. 2). The DLP contains an omnidirectional microphone with a flat 20-20 kHz frequency response. Frequencies above 10 kHz are suppressed, as they are unlikely to contain human speech activity. Low frequency sound is suppressed by a 70 Hz high-pass filter. Digital data were recorded using a 10 kHz low-pass filter to suppress high-frequency sounds. Frequencies were recorded using a 16 kHz 16-bit sigma-delta analog to digital (ADC) converter with 8x over-sampling digital interpolation.

The LENA DLP has a hardware limiter for intensity suppression at 87dbSPL. This limiter was not problematic for this study, given that normal conversational speech averaged 81 dB-SPL at the microphone-to-mouth distance used of 9 inches. No clipping was observed in the sentence samples that were gathered. When tested on various frequencies, the DLP showed an equal flat response at 125 Hz, 250 Hz,

500 Hz, 750 Hz, 1000 Hz, 1500 Hz, 2000 Hz, 3000 Hz, 4000 Hz, 6000 Hz, and 8000 Hz.

The DLP was calibrated using 500Hz calibration tones at 65, 70, 75, 80 and 85 dB-SPL. The tones were delivered via Grason-Stadler GSI 61 clinical audiometer and factory Grason-Stadler speakers in a sound attenuated room. The DLP and a sound level meter (set at 50 dB low, C-weight, and fast speed) were both 9 inches from the speaker. These calibration recordings on the DLP were saved for linear interpolation to convert into dB-SPL from Praat (Boersma & Weenink, 2010).

The slopes and y-intercepts derived from the linear interpolation for each of the two DLPs used for the study were: (DLP 1, slope=1.0447, y-intercept=9.146,  $R^2=.9997$ ; DLP 2, slope=1.0447, y-intercept=12.784,  $R^2=.9997$ ).

The DLP was clipped on participants' chests 9 inches below their mouth, to record spontaneous speech and the assessment tasks. Recordings from the natural environment were made in the participants' home and other non-clinical settings in which their primary conversational partner was usually their spouse, with informed consent. Participants were instructed on maintaining a microphone-to-mouth distance and given a ribbon to maintain the same mouth-to-microphone distance over the two days of recording. They were also provided with written instructions on making recordings with the DLP. None of the participants were wearing the SpeechVive™ during speech recordings, prior to the 12 weeks period or after device use for 12 weeks.

### **Speech Tasks**

Experimental sessions began with reviewing consent, fitting of the LENA-DLP, and casual conversation. The Voice Handicap Index (VHI) (Jacobson et al., 1997; C. A. Rosen et al., 2004) and Beck Depression Inventory (BDI) (Beck et al., 1996) were collected. After approximately 15 minutes of conversation, participants were given a hearing screening and the MMSE (Crum, Anthony, Bassett, & Folstein, 1993; Folstein et al., 1975). Instructions were given for formal speech testing, which included the Sentence Intelligibility Test (Yorkston et al., 1996) and a script recording of speech tasks, included for a previous study.

The participants were then instructed on the LENA-DLP, recording instructions, and confirming microphone-to-mouth distance of 9 inches. Participants were given written instructions for device use and a letter to inform any other conversational partners that they may encounter during the study about the recording. They were instructed on how to switch the device on and off should they or their conversational partner wish to have a private conversation. Participants then completed 15 hours of recording within their home environments over two days. They were instructed not to use the DLP in the presence of others not made aware that the device was recording and agreeing to have their speech recorded.

After recording in the home environment over two days, the participants immediately enrolled in the SpeechVive™ treatment study. During this time, they wore the SpeechVive™ device for 3-8 hours each day for 12 weeks. The same experimenter visited each participant every 2 weeks in their home to gather data for the treatment study. At the end of the 12-week SpeechVive™ study, the home study

was repeated while wearing and not wearing the device. After completing the post-SpeechVive™ recording, and returning the device to the examiner, each participant received the LENA-DLP and repeated the study within 24 hours or less of last wearing the SpeechVive™ device. A 16-hour sample was taken from the home environment including spontaneous conversation with the primary communication partner.

### **Data Processing and Analysis**

#### **Outcome Variable: Intelligibility**

The same procedures as reported in Rountrey et al., in preparation, were used for this pilot study. Spontaneous utterances from the home environment were reduced to conversational segments (periods without voice removed), time stamped, and selected via a random number generator. Spontaneous utterances were selected from the sound file by matching the random number generator to timestamps on the recording. The nearest sentence within a conversation that met the following criteria was selected: From intentional 1-on-1, person-to-person communication, between 4 and 15 words long without background noise interfering with speech, and a unique sentence. Sentences were excluded if the conversation was with more than one person, the conversational partner was not aware they were being recorded, sentences were too long such that listeners' memory would be taxed, or were repeated sentences. Ten randomly selected sentences from the home environment were saved in a digital audio file (.wav) for intelligibility analysis.

Three different novel, inexperienced listeners transcribed each speaker's sentences, with no one listener hearing the same sentence more than once. Each



volunteer listener passed a hearing screening, and was without delay, disorder, or disease that would affect hearing and transcribing the sentences. Each was limited to 20 minutes of listening to reduce listener fatigue. Listeners used laptops to transcribe what they heard as the sentences were played for them. Amplification levels on the speakers were set to match the same sound level meter values as were read during the calibration tone session. Listeners were instructed that they would hear each sentence only once and that they were to transcribe exactly what they heard, without regard for punctuation or capitalization. They were told not to make any corrections until the entire sentence was transcribed, to reduce their memory load. Pauses were given until the listener was ready to start the next sentence and after each set of 10 sentences.

Sentences were transcribed into Microsoft Excel spreadsheets and compared with transcriptions made by the researchers, who listened to the sentences in context for an unlimited number of times and, when needed, at a higher intensity.

Comparisons were made with R, Package 'qualV' (van den Boogaart et al., 2014) using the Least Common Sequence (LCS) function algorithm. Prior to running the program, the entire spreadsheet was checked for spelling errors. The LCS algorithm was adapted to do a number of things: text was converted to lowercase without punctuation, single spaces between words and delete leading spaces, so as not to count these as transcriber errors. The algorithm then created output columns that look for the least common sequence, or the number of identical words that are in the same sequence as the key sentence. Then, the number of identical words in the sequence was divided by the greatest number of words in either sentence, so

that errors of omission and commission are accounted for. This percentage was used for “intelligibility.”

To determine the accuracy of the LCS algorithm, a random sample of 10 sentences were calculated for intelligibility by hand, then compared to the LCS code written for R, resulting in a perfect correlation ( $R^2=1.0$ , 0% error). The percent intelligibility based on R was determined for each of the participants’ in each environment.

### **Outcome Variable: intensity**

Intensity of the utterance was measured using Praat (Boersma & Weenink, 2010) software for speech analysis. Praat intensity was set to show a range of 25dB to 100dB with “subtract mean pressure” checked as OFF. If pauses in the midst of an utterance were greater than 150 ms, they were deleted (Hammen & Yorkston, 1996) before the mean intensity was computed, from the beginning to the end of the sound contour. The intensity level was converted into decibels in sound pressure level re .0002 dynes/cm<sup>2</sup> using the slope and y-intercept of the linear interpolation derived from measurement of calibration tones. The mean of each participant’s sentences in each environment was calculated and saved as “intensity” for statistical analysis.

### **Statistical Methods**

Repeated measures ANOVAs were used compare pre- and post- treatment values of percent intelligibility and intensity in spontaneous speech.

## **Results**

### **Participants**

Seven participants with idiopathic Parkinson's disease (5 men, 2 women, 57- to 74 years, median 70 years) participated in this study (**Table 1**). All participants were on drug therapy and earned a minimum score of 23/30 on the MMSE (Folstein et al., 1975). The range of overall impairment at the time of speech evaluation extended from mild to moderately severe disability. Two participants had deep brain stimulators and settings remained constant during the study. Of the 7 participants, 4 reported no history of speech therapy, 2 reported Lee Silverman Speech Therapy (LSVT), or "4 weeks of talking loud," and 1 reported another type of speech therapy, the most recent speech therapy being 1.5 years prior to participation in the study. There were a total of 10 volunteers who participated in the pre-SpeechVive™ recording and volunteered for the SpeechVive™ study. Two did not complete 12 weeks with the SpeechVive™, and one opted out of the post-SpeechVive™ recording.

### **Listeners**

Fifty-eight total listeners participated in the study, the same listeners who participated in Rountrey et al., in preparation. Their ages ranged from 18 to 53 years (mean 21.9 years). None of the listeners had experience or completed coursework in motor speech disorders or voice at the time of their participation. All listeners were native speakers of Standard American English, free of neurological or psychiatric disorders, and with visual acuity adequate to perform the task, and passed a hearing screening. Sixty volunteered, one was excluded due to not passing the hearing screening, and one was excluded after self-reporting an auditory

processing disorder. Most listeners did one session per day ( $n=53$ ), and no more than three sessions per day ( $n=3$ ).

### **Repeated Measures ANOVA**

SPSS version 21 was used for statistical analysis (SPSS, 2012). Repeated measures ANOVAs with a Greenhouse-Geisser correction determined over time, before and after application of SpeechVive™ device for 12 weeks. There was no significant difference in intelligibility  $F(1,6)=.176$ ,  $p=.689$ , or intensity  $F(1,6)=.021$ ,  $p=.890$ . Therefore, it is concluded that, speech that is unaided does not show changes following application of the SpeechVive™ device for 12 weeks.

### **Power Analysis**

A power analysis was conducted using G\*Power version 3.1.3 (Faul, Erdfelder, Lang, & Buchner, 2007), using the data of these 7 participants. Percent intelligibility change over the treatment period effect size was negligible, Cohen's  $d = .157$ . Change in intensity was also negligible, Cohen's  $d = .055$ . The numbers of participants needed were also computed for a power of .80. For detecting change over the treatment period in percent intelligibility, 314 participants would be required. For detecting change over the treatment period in intensity, 2597 participants would be required.

### **Discussion**

Application of the SpeechVive™ device for 12 weeks did not elicit a difference in unaided spontaneous speech intelligibility or intensity in the home, when comparing pre-treatment and post-treatment within 24 hours of cessation of use of the SpeechVive™. Further investigation with a larger sample size is not warranted

given the small effect size when the patients were unaided. However, the sample was comprised of only mild to moderately affected patients who may be less susceptible to treatment effect, and repeating a similar experiment with a wide range of impairment may yield different results.

In the present study, recordings were made anywhere the participant went in the house, and participants were engaged in other activities besides pure speech tasks. Differences in cognitive load and competition for cognitive resources have been shown to have performance effects and distract attention from speech (Bunton & Keintz, 2006; Dromey & Benson, 2003).

Given our findings of limited effects on unaided speech, the effect of the aid may only be evident when the aid is being worn and turned on. However, six of the participants in this study were judged to have mildly to moderately affected speech/voice, while only one was moderate-to-severely affected. The patient with the most severe impairment saw the greatest gain in intelligibility (difference pre-post=17.58%). This may indicate that the treatment might have a greater effect on those with the more severe intelligibility problems. Expanding the number of participants for this study with more severe speech impairments is recommended for future study. However, the feasibility of a 12-13 week study with 320 participants or more should be considered.

Methods of investigating functional gains in home and natural environments are in development, from dosimetry (Hillman, Heaton, Masaki, Zeitels, & Cheyne, 2006b) to use of accelerometers (Schloneger & Hunter, 2015) to encrypted digital recorders (Ford et al., 2008). In addition many remote sensing devices with

analysis will become available with smartphones (Mehta et al., 2012). Exercises that are carried on at home will need to be monitored for efficacy in the functional environment of the patient, rather than in the clinic.

It is also important to note that recording large spontaneous speech samples from the home environment is a unique and developing methodology in and of itself. As technology advances, efficient and effective measurement of functional speech in the home and other natural environments of those with voice and speech difficulties should continue to progress. Practitioners are in need of normative values for acoustics and intelligibility as measured in the home environment for healthy controls as well as persons with PD and other disorders affecting communication. Some of this information may be available in the data corpus from this study. Continuing to analyze data from additional patients and healthy controls could yield additional information on spontaneous speech in the home.

### **Contributors**

CER participated in the original initiation and design of the study, devised procedures, took part in recruitment, data collection, processing and analysis. She was the principal investigator and the lead author in preparing drafts.

NMB participated in data collection and processing, lab organization, and commenting on draft versions of the work.

JH, through Purdue University, provided equipment and participant honorariums for the SpeechVive™ study, which ran in tandem with this study, reviewed and commented on draft versions of the work.

CLL participated in the original initiation and design, data analysis, and reviewing and critically commenting on drafts and the final version.

### **Acknowledgements**

The authors gratefully acknowledge funding from James Madison University department of Communication Sciences and Disorders, which supported CER as a Ph.D. candidate and NMB as a master's student during this time. Purdue University provided funding to pay participants during the SpeechVive™ study, which ran in tandem with this study. Purdue University also paid mileage to CER for visits to patients' homes during the treatment study.

### **Disclosures**

JH has financial interest in SpeechVive™, Inc. and sits on the Board of Directors for the Company.

# References

- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio*,
- Boersma, P., & Weenink, D. (2010). {P} raat: Doing phonetics by computer.
- Bunton, K., & Keintz, C. (2006). Effects of a concurrent motor task on speech intelligibility for speakers with parkinson disease. Paper presented at the *Conference on Motor Speech, Austin, TX*,
- Countryman, S., & Ramig, L. (1993). Effects of intensive voice therapy on voice deficits associated with bilateral thalamotomy in parkinson's disease: A case study. *Journal of Medical Speech-Language Pathology, 1*(4), 233-250.
- Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population-based norms for the mini-mental state examination by age and educational level. *Jama, 269*(18), 2386-2391.
- Dromey, C., & Benson, A. (2003). Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research, 46*(5), 1234-1246.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G\* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198.
- Ford, M., Baer, C. T., Xu, D., Yapanel, U., & Gray, S. (2008). *The LENATM Language Environment Analysis System: Audio Specifications of the DLP-0121*,



- Forrest, K., Weismer, G., & Turner, G. S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults. *The Journal of the Acoustical Society of America*, 85(6), 2608-2622.
- Fox, C. M., Morrison, C. E., Ramig, L. O., & Sapis, S. (2002). Current perspectives on the lee silverman voice treatment (LSVT) for individuals with idiopathic parkinson disease. *American Journal of Speech Language Pathology*, 11(2), 111-123.
- Hammen, V. L., & Yorkston, K. M. (1996). Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria. *Journal of Communication Disorders*, 29(6), 429-445.
- Hillman, R. E., Heaton, J. T., Masaki, A., Zeitels, S. M., & Cheyne, H. A. (2006). Ambulatory monitoring of disordered voices. *The Annals of Otology, Rhinology, and Laryngology*, 115(11), 795-801.
- Ho, A. K., Iannsek, R., & Bradshaw, J. L. (1999). Regulation of parkinsonia speech volume: The effect of interlocuter distance. *Journal of Neurology, Neurosurgery, and Psychiatry*, 67, 199-202.
- Ho, A. K., Bradshaw, J. L., Iannsek, R., & Alfredson, R. (1999). Speech volume regulation in parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia*, 37(13), 1453-1460.
- Jacobson, B. H., A., J., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI): Development and validation. *American Journal of Speech Language Pathology*, 6, 66-70.
- Mehta, D. D., Zanartu, M., Feng, S. W., Cheyne, H. A., & Hillman, R. E. (2012). Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *Biomedical Engineering, IEEE Transactions On*, 59(11), 3090-3096.
- Morris, M. E., Iannsek, R., Matyas, T. A., & Summers, J. J. (1994). The pathogenesis of gait hypokinesia in parkinson's disease. *Brain*, 117, 1169-1181.

- Oliveira, R. M., Gurd, J. M., Nixion, P., Marshall, J. C., & Passingham, R. E. (1997). Micrographia in parkinson's disease: The effect of providing external cues. *Journal of Neurology, Neurosurgery, and Psychiatry*, 63, 429-433.
- Ramig, L. O., Countryman, S., Thompson, L. L., & Horii, Y. (1995). Comparison of two forms of intensive speech treatment for parkinson disease. *Journal of Speech, Language, and Hearing Research*, 38(6), 1232-1251.
- Ramig, L. O., & Dromey, C. (1996). Aerodynamic mechanisms underlying treatment-related changes in vocal intensity in patients with parkinson disease. *Journal of Speech, Language, and Hearing Research*, 39(4), 798-807.
- Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. *The Laryngoscope*, 114(9), 1549-1556.
- Sadagopan, N., & Huber, J. E. (2007). Effects of loudness cues on respiration in individuals with parkinson's disease. *Movement Disorders*, 22, 651-659.
- Schloneger, M., & Hunter, E. J. (2015). Correlations between simultaneous voice qualify readings by accelerometer and acoustic transducers. *11th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research*, London, United Kingdom.
- SPSS, I. (2012). IBM SPSS statistics version 21. *Boston, Mass: International Business Machines Corp*,
- Taaffe, D. R., Duret, C., Wheeler, S., & Marcus, R. (1999). Once-weekly resistance exercise improves muscle strength and neuromuscular performance in older adults. *Journal of the American Geriatrics Society*, 47(10), 1208-1214.
- van den Boogaart, K., Rost, S., Petzoldt, T., & Petzoldt, M. T. (2014). Package ‘qualV’.
- Yorkston, K. M., Beukelman, D. R., & Tice, R. (1996). Sentence intelligibility test. *Tice Technology Services*,

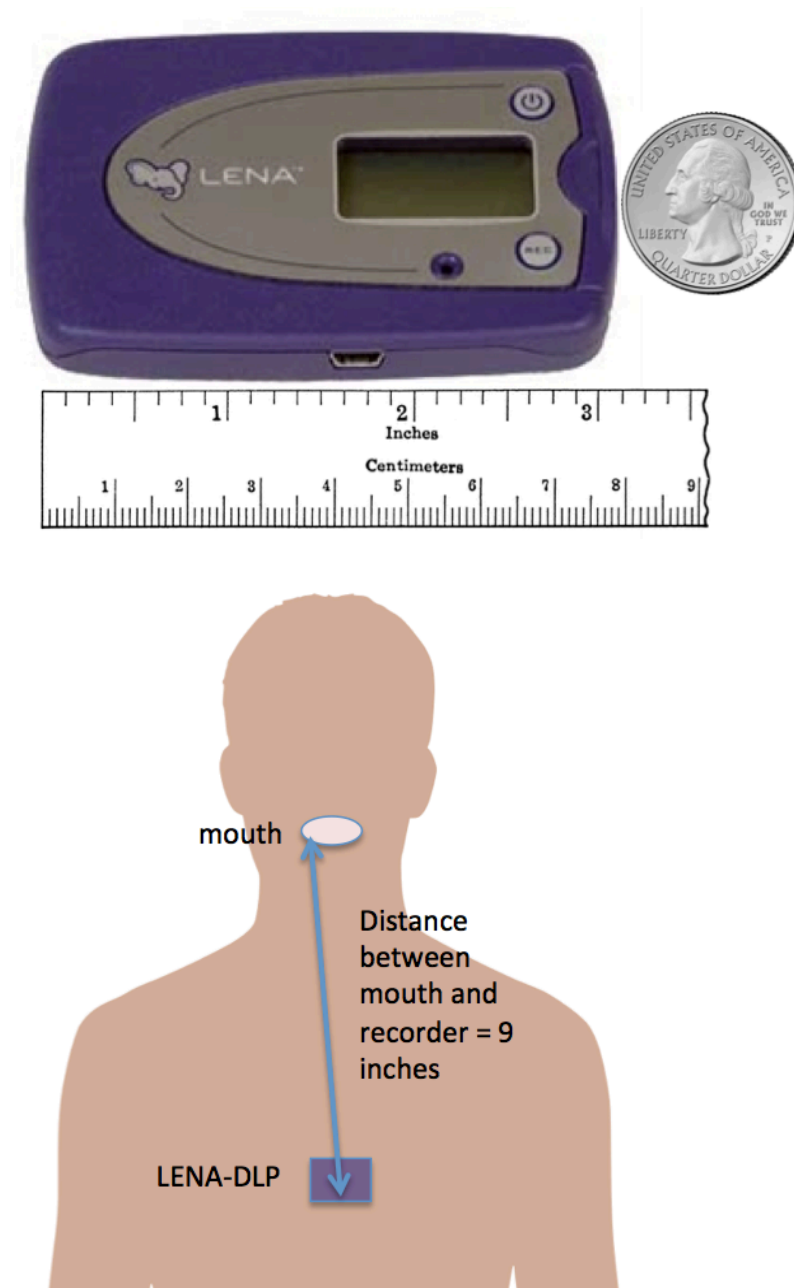
***Table 1: Participants with Parkinson's Disease Clinical Data (n=7)***

ID	Age	Gender	Voice	Beck	Years	MMSE
			Handicap Index	Depression Index	since Diagnosis	
301	72	F	50	19	10	27
302	74	M	40	3	10	28
306	70	M	37	0	13	29
307	63	F	52	0	10	27
308	65	M	60	17	15	30
309	57	M	8	9	2	29
311	73	M	18	2	2	29

***Figure 1: The SpeechVive™ wearable device for treatment of hypophonia in Parkinson's Disease***



**Figure 2: LENA-DLP.** The DLP is 2.5oz, measures approximately 8cm x 5cm x 1cm. Shown here compared to a U.S. quarter (2.4cm), and a schematic measuring the distance between mouth and microphone on DLP, as participants were instructed.



## Conclusions

This project set out to quantify the speech and voice behavior of persons with Parkinson's disease in their home environments, contrast that environment with the clinical/laboratory environment, assess the predictability of home speech behavior from clinical tasks, and measure changes in home speech behavior after treatment. It was hypothesized that environment effects would be significant, that significant clinical predictors would emerge, and that therapy would change speech behavior over time. Many of the results of this inquiry were surprising.

It was concluded that when it comes to the "hallmark" sign of hypophonia (reduced vocal loudness) in Parkinson's disease, the participants with PD who enrolled in this study did not differ from the healthy controls on speech intensity of spontaneous sentences in either the home or clinical environments. Reduced intelligibility was evident in our participants rather than reduced intensity, regardless of environment. When we investigated common components of a clinical evaluation, we found that vocal intensity on a sustained /a/ was only useful as a predictive measure of home intensity when it was combined with the Sentence Intelligibility Test (SIT) (Yorkston et al., 1996), and that the SIT were significantly predictive of home intelligibility on their own. When looking at treatment, we found no significant gains in home spontaneous speech intensity or intelligibility after 12 weeks of SpeechVive™ use, however our study was limited by a small, rather homogenous sample group and only examined them when not wearing the device.

These studies together demonstrate a renewed importance of attending to differences between environment and task when assessing functional speech, and

that clinical measures are not necessarily accurate for predicting speech behavior at home. New technologies are emerging such as those used for this study and others that may have smartphone platforms and increased power and portability.

Ambulatory measures of speech and remote sensing of speech behavior and the functional environment is a growing field with much investigation needed.

Standardized measures and efficient and effective methodologies are in development for truly functional measures of speech behavior in the home for baselines and measurement of home therapy activities and progress.