### A LONGITUDINAL STUDY OF PARKINSONIAN SPEECH CHARACTERISTICS

By

Juliane Leigh Brinkman

### A THESIS

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Communicative Sciences and Disorders- Master of Arts

#### ABSTRACT

# A LONGITUDINAL STUDY OF PARKINSONIAN SPEECH CHARACTERISTICS By

#### Juliane Leigh Brinkman

It is well documented that changes in speech occur with age and can be augmented by the presence of neuropathology (Little et. al, 2009; Ramig & Ringel, 1983; Sadagopan & Smith, 2013). Parkinson's disease (PD) is a progressive, neurodegenerative disorder that can have a profound effect on speech and its subsystems. Advances in technology and the use of acoustic applications have made it possible to objectively measure PD speech, and the current literature reports success in accurately differentiating PD speakers from healthy speakers. The present study was conducted with two goals in mind. First, to determine whether acoustic variables can be used to differentiate PD speakers from healthy speakers during times before the clinical diagnosis of PD. Second, to determine the usefulness of acoustic variables for monitoring the disease progression of PD. To achieve this, recordings spanning two to seven decades were obtained for 10 PD speakers and 9 control speakers. Recordings were edited for target speaker and assigned to a speaker group: PD Pre-Dx, PD Post-Dx, or Control. Values for voice and articulation measures were obtained from the recordings, and ANOVA and ANCOVA analyses were completed using MATLAB. A second ANOVA and ANCOVA analysis was performed on a combined PD speaker group and the Control group. Several acoustic measures demonstrated PD Pre-Dx -group mean values significantly less than Control group mean values, making them a potential tool for early PD detection. Other acoustic measures demonstrated trends that may show potential in future research for use in PD progression monitoring. Limitations of study and suggestions for future research are discussed.

Copyright by Juliane Leigh Brinkman 2014 To my family. It's impossible to measure the amount of unconditional love and encouragement you all have given me.

#### ACKNOWLEDGMENTS

I would like to acknowledge my committee members for their immense support of my academic endeavors. Thank you, Dr. Shrivastav, for the initial idea that allowed for the formation of this thesis project, and your support and positive attitude along the way. Thank you, Mark Skowronski for your instruction in the lab during the research process, enthusiasm for science, and encouragement, especially on those multiple Saturday meetings. Your patience is commendable! Thank you, Dr. LaPine, for your tremendous influence in the classroom and for your insight as my advisor and committee member. Also, thanks for the laughs along the way. In fact, I think you still owe me a bottle of wine...

I am delighted to have had the opportunity to share the lab with such admirable colleagues. Supraja, thank you for your guidance during the early stages of my project and your help in recruiting participants. Lisa, thank you for your veteran insight regarding the research process, as well as your help and support in the lab. Deanna, thank you for enduring this process with me- we made it!

A special thanks to the Vincent Voice Library at Michigan State University for their assistance in providing me with the recordings I needed to complete this project.

Finally, I would like to recognize the immeasurable support that my professors, clinical supervisors, colleagues, family, and friends have given me throughout my graduate studies. I cannot thank you all enough!

v

# TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
KEY TO ABBREVIATIONS	xii
INTRODUCTION	1
Background	1
Disease Progression	2
Diagnosis	4
Treatment	4
Pharmacological Treatment Options	4
Surgical Treatment Options	5
Hypokinetic Dysarthria	6
Treatment for Hypokinetic Dysarthria	7
Acoustic Applications for PD	8
Digitizing Sound for Acoustic Applications	9
Spectral-Based Analysis	11
Recording Speech for Analysis	13
Acoustic Measures for PD	14
Voice Quality Measures	14
Articulation Measures	18
The Present Study	20
METHODOLOGY	22
Design Type	22
Participants	22
Data Collection	22
Procedure	27
RESULTS	
PT mean ST: 1-way ANCOVA with Post Hoc Testing	29
PT mean ST: 1-way ANOVA with Post Hoc Testing	
PT std ST: 1-way ANCOVA with Post Hoc Testing	
PT std ST: 1-way ANOVA with Post Hoc Testing	
PS mean: 1-way ANCOVA with Post Hoc Testing	34
PS mean: 1-way ANOVA with Post Hoc Testing	35
PS mean: 1-Way ANOVA	45
CPP: 1-way ANCOVA with Post Hoc Testing	36
CPP: 1-way ANOVA with Post Hoc Testing	
CC SDS: 1-way ANCOVA with Post Hoc Testing	

CC SDS: 1-way ANOVA with Post Hoc Testing	39
DCC SDS: 1-way ANCOVA with Post Hoc Testing	40
DCC SDS: 1-way ANOVA with Post Hoc Testing	41
Combined PD Talker Group: 1-way ANCOVA with Post Hoc Testing	42
Combined PD Talker Group: 1-way ANOVA with Post Hoc Testing	43
DISCUSSION	46
Differences in Group Means Determined by ANOVA	46
Differences in Regression Lines Determined by ANCOVA	47
Implications of Study Findings	47
Limitations	49
Suggestions for Future Research	49
APPENDIX	50
REFERENCES	52

### LIST OF TABLES

Table 1: Sources Used to Obtain Recordings
Table 2: Types of Recordings Used in Analysis and Time Range for PD Speakers25
Table 3: Types of Recordings Used in Analysis and Time Range for ControlSpeakers
Table 4: ANCOVA Test Results- PT mean ST
Table 5: ANOVA Test Results- PT mean ST
Table 6: ANCOVA Test Results- PT std ST
Table 7: ANOVA Test Results- PT std ST
Table 8: ANCOVA Test Results- PS mean
Table 9: ANOVA Test Results- PS mean
Table 10: ANCOVA Test Results- CPP
Table 11: ANOVA Test Results- CPP
Table 12: ANCOVA Test Results- CC SDS
Table 13: ANOVA Test Results- CC SDS
Table 14: ANCOVA Test Results- DCC SDS
Table 15: ANOVA Test Results- DCC SDS
Table 16: ANCOVA Test Results- CPP
Table 17: ANOVA Test Results- PT mean ST44
Table 18: ANOVA Test Results- CC SDS
Table 19: ANOVA Test Results- DCC SDS
Table 20: Type of Recording and Year for Individual Speakers

## LIST OF FIGURES

Figure 1: Frequency in Semitones compared to Frequency in Hz17
Figure 2: Time Range of Recordings and Time Points of Recordings Used for Analysis- PD Participants
Figure 3: Time Range of Recordings and Time Points of Recordings Used for Analysis- Control Participants
Figure 4: Regression Line Slopes- PT mean ST
Figure 5: ANCOVA Post Hoc Test- PT mean ST
Figure 6: ANOVA Post Hoc Test- PT mean ST
Figure 7: Regression Line Slopes- PT std ST
Figure 8: ANCOVA Post Hoc Test- PT std ST
Figure 9: ANOVA Post Hoc Test- PT std ST
Figure 10: Regression Line Slopes- PS mean
Figure 11: ANCOVA Post Hoc Test- PS mean
Figure 12: ANOVA Post Hoc Test- PS mean
Figure 13: Regression Line Slopes- CPP
Figure 14: ANCOVA Post Hoc Test- CPP
Figure 15: ANOVA Post Hoc Test- CPP
Figure 16: Regression Line Slopes- CC SDS
Figure 17: ANCOVA Post Hoc Test- CC SDS
Figure 18: ANOVA Post Hoc Test- CC SDS
Figure 19: Regression Line Slopes- DCC SDS
Figure 20: ANCOVA Post Hoc Test- DCC SDS

Figure 21: ANOVA Post Hoc Test- CC SDS	41
Figure 22: Combined PD Group Regression Line Slopes- CPP	42
Figure 23: ANCOVA Post Hoc Test- CCP	43
Figure 24: Combined PD Group ANOVA Post Hoc Test- PT mean ST	44
Figure 25: Combined PD Group ANOVA Post Hoc Test- CC SDS	44
Figure 26: Combined PD Group ANOVA Post Hoc Test- DCC SDS	45

### KEY TO ABBREVIATIONS

 $\Delta$ HFCC: delta human factor cepstral coefficient

ANCOVA: analysis of covariance

ANOVA: analysis of variance

ANSI: American National Standards Institute

AUD-SWIPE: Auditory-Based Spectrum-Enhancing Preprocessing Stage Sawtooth Waveform Inspired Pitch Estimator

CPP: cepstral peak prominence

CC SDS: cepstral coefficient standard deviation sum

DAF: delayed auditory feedback

DCC SDS: delta cepstral coefficient standard deviation sum

EMST: expiratory muscle strength training

HFCC: human factor cepstral coefficient

Hz: Hertz

LB: Lewy Body

LN: Lewy Neurite

LSVT: Lee Silverman Voice Treatment

MFCC: mel factor cepstral coefficient

PD: Parkinson's disease

PD Post-Dx: Parkinson's disease group, post-diagnosis

PD Pre-Dx: Parkinson's disease group, pre-diagnosis

PT mean ST: pitch mean, semitones

PT std ST: pitch standard deviation, semitones

PS mean: pitch strength mean

RLS: regression line slope

#### INTRODUCTION

#### Background

It is well documented in the current literature that changes in speech production related to aging occur throughout one's lifetime and can augmented by the presence of neuropathology (Little et. al, 2009; Ramig & Ringel, 1983; Sadagopan & Smith, 2013). Parkinson's disease (PD) is one such neuropathy which exacerbates the age-related changes in speech, as it is a neurodegenerative disease of the central nervous system that primarily disturbs motor function.

PD is the second most common neurodegenerative disorder with a prevalence of 1-2 cases per 1,000 and 10 cases in 1,000 in the population that is 65 years of age and older (Factor & Weiner, 2008; McNeil, 2009). The average age of onset of PD is 60 years; however, 10% of patients show symptoms before the age of 40 years (McNeil, 2009). Researchers estimate that there were approximately four million people over the age of 50 with PD worldwide in 2005, and predict that this figure will increase to more than 8 and a half million by the year 2030 (Dorsey et al., 2007).

PD is exceedingly expensive to manage due to its disease characteristics and slow progression. In the United States, nearly \$25 billion is spent every year on the direct and indirect costs of PD (Parkinson's Disease Foundation, 2014). These expenses include, but are not limited to hospital and specialized clinic visits, disability benefits, and other medical expenses such as home equipment installation and caregiver costs. Pharmaceutical treatment costs average \$2,500 annually for individuals with PD, and surgery to manage severe symptoms can cost up to

\$100,000 per procedure (Parkinson's Disease Foundation, 2014). PD can also have a profound effect on the individual's quality of life as well as family and social dynamics.

#### **Disease Progression**

PD is distinguished by distinctive lesions in nerve cells within the brain that are susceptible to the disease. These lesions are called Lewy Neurites (LNs) and Lewy Bodies (LBs) (Braak et al., 2003). It is not yet known what causes the spindle-shaped LNs and spherically-shaped LBs to form, thus the cause of PD is a mystery. Many genetic studies have been conducted with the hope of discovering the cause of PD and all yield fairly inconclusive results. It is currently believed that the development of PD involves a genetic susceptibility in combination with certain undetermined environmental factors (McNeil, 2009).

The nerve cells that are vulnerable to LNs and LBs experience pathological changes at different times during the progression of PD. PD progression can span over several decades and rate of progression is different for each individual (Parkinson's Disease Foundation, 2014). It is standard for the progression of PD to be divided into six stages which are related to the order in which LNs and LBs appear in brain structures (Braak et al., 2004). The first three stages are considered to be the pre-symptomatic stages since lesions begin developing a significant amount of time before the appearance of overt motor dysfunction. The final three stages are considered to be the post-symptomatic stages since visible

 $\mathbf{2}$ 

symptoms are apparent. These stages are not only useful for describing the disease's progression, but also for making differential diagnosis.

The progression of PD pathology begins at the level of the brainstem and is nearly confined to the medulla oblongata until further upwards progression occurs during the mid-stages of the disease (Braak et al., 2003). Symptoms commonly reported during the early stages of PD progression include changes in the senses of smell and taste. Early symptoms related to major visceral functions include disturbed motor innervation of the heart, lungs, pharynx, and esophagus (Hawkes & Deeb, 2006). Additional functions involved at this time include coughing, gagging, vomiting, swallowing, and phonation (Hawkes & Deeb, 2006). As the disease progresses upwards throughout the brainstem, blood pressure, sleep, and inhibition of somatosensory and visceral pain input are all affected (Braak et al., 2003; Hawkes & Deeb, 2006).

As the middle stages of PD emerge, the disease process continues to progress upwards and the first solitary LNs can be seen in the pars compacta of the substantia nigra. Motor symptoms will become visually obvious and a diagnosis of PD can be made at this point (Braak et al., 2004). Symptoms observed during the middle stages of PD progression involve the functions of voluntary movement, sleepwake regulation, arousal, learning, and memory (Hawkes & Deeb, 2006).

During the final stages, the disease process will continue to move upwards into cortical regions as dopaminergic neurons continue to be depleted in the substantia nigra. Cortical damage disturbs the functions of memory and emotion

(Braak et al., 2004; Hawkes & Deeb, 2006). Patients will experience a decline in most cognitive functions and manifest in the full range of PD-associated clinical symptoms at this time (Hawkes & Deeb, 2006; Jankovic, 2008).

#### Diagnosis

The cardinal features of PD are tremor at rest, rigidity, akinesia, and postural instability (Jankovic, 2008). Currently, clinical diagnosis of PD is made by most experts when two of the cardinal features are present and a consistent, positive response to levodopa pharmaceutical treatment is seen in motor function (Bartels & Leenders, 2009). Despite the fact that the disease progression is well classified, a true diagnosis of PD can only be confirmed by a post-mortem examination of brain tissues for LBs and LNs.

#### Treatment

Pharmacological Treatment Options. Presently, there is no cure for PD available. PD management options are palliative and offer patients a means of improving quality of life. In a review of the literature conducted by Schulz and Grant (2000), three primary treatments for PD were identified: pharmacological treatments, surgical treatments, and fetal cell transplantation. Many PD patients receive pharmacological treatments to manage motor dysfunction. Pharmacological treatments aim to replace or enhance dopamine present in the brain affected by PD. Medications that enhance dopamine are dopamine agonists, and medications that replace dopamine involve levodopa, usually combined with carbidopa (Schulz & Grant, 2000). Patients taking levodopa pharmaceuticals for the treatment of PD

experience what is termed "on and off" effects of the medications. Patients will see the greatest reduction of motor symptoms soon after a dose is taken, but these effects will diminish as time elapses. The results of levodopa medication become shorter in duration after many years of using the drug (Schulz & Grant, 2000). Thus, the patient will endure longer "off" periods as the disease progresses in which motor dysfunction will be more severe.

Surgical Treatment Options. Surgical interventions are also available for the treatment of PD. Surgical options include ablative surgeries, deep brain stimulation (DBS), and restorative procedures. Ablative surgeries are thalamotomy, which involves creating lesions in the thalamus, and pallidotomy, which involves creating lesions in the globus pallidus of the basal ganglia (Schulz & Grant, 2000). Thalamotomy reduces contralateral tremor and rigidity, and pallidotomy reduces the effects of all major PD symptoms (Schulz & Grant, 2000).

DBS is often performed in place of ablative surgeries because it has a lower risk of causing permanent neurological deficits. DBS is the electrical stimulation of the thalamus, subthalamic nucleus, or globus pallidus internal to reduce PD symptoms (Schulz & Grant, 2000). Symptom reduction is dependent on the brain structure involved in the DBS. Stimulation of the thalamus is effective in improving essential and resting tremor. Stimulation of the subthalamic nucleus improves rigidity and akinesia. Stimulation of the globus pallidus internal improves dyskinesias and rigidity, but not as profoundly as stimulation of the subthalamic nucleus (Schulz & Grant, 2000).

 $\mathbf{5}$ 

Fetal cell transplantation is a restorative procedure that is still considered to be experimental. During this procedure, fetal dopaminergic cells are transplanted in the caudate or putamen of the basal ganglia of the brain affected by PD (Schulz & Grant, 2000). Improvements in rigidity, hypokinesia, and effects of levodopa were observed for experimental trials (Schulz & Grant, 2000).

#### Hypokinetic Dysarthria

In addition to the aforementioned symptoms of PD, speech and its subsystems of respiration, phonation, articulation, resonance, and prosody are also significantly affected during disease progression. Blocket and colleagues (2011) report that that subsystems of speech most affected by PD, in order, are phonation, articulation, and prosody. The degree of dysfunction and severity in the subsystems of speech vary in individuals with PD (Ramig et al., 2008). Speech and speechrelated symptoms are often overlooked as it is estimated that 89% of people with PD have some type of speech and/or voice disorder, but only 3-4% of these people receive speech and voice treatment (Ramig et al., 2008).

Hypokinetic dysarthria is the name for the speech production disorder resulting from PD. A classic study conducted by Darley, Aronson, and Brown (1969) distinguished the prominent features of hypokinetic dysarthria. Breathy voice, harsh voice, and low pitch are qualities of PD speech and are thought to be related to the rigidity of laryngeal musculature (Darley et. al, 1969). PD speech is also characterized by prosodic insufficiency which includes mono-loudness, reduced stress, short rushes of speech, variable rate, and imprecise consonants, which can

be attributed to the fast repetitive movements of the reduced articulatory range seen in PD (Darley et. al, 1969). These changes decrease the speaker's intelligibility and efficiency of communication. In turn, the speech changes caused by hypokinetic dysarthria could have substantial influence on one's quality of life. For instance, changes may occur in one's social activities and identity, and in some cases social isolation can transpire.

Treatment for Hypokinetic Dysarthria. Unfortunately, pharmaceutical and surgical treatments tend to have little impact on the speech and voice impairments associated with PD. However, some improvements in communication have been demonstrated by PD patients who have received traditional speech therapy. Speech intervention for people with PD typically involves the teaching and implementation of strategies to compensate for the features of hypokinetic dysarthria. The most commonly marketed speech intervention used for people with PD is the Lee Silverman Voice Treatment (LSVT) protocol. The LSVT involves a rigorous treatment program to increase the loudness of people with PD to improve their overall communication abilities.

Speech-language pathologists who are not officially trained in the LSVT protocol incorporate a variety of compensatory strategies into speech therapy regimens. Following a thorough assessment, the therapist may choose to focus on one or several aspects of speech production of a client with PD. A comprehensive approach likely contains elements to address loudness, breath support, variation in pitch, speaking rate, intonation, intelligibility, and articulation precision, speed,

and strength (Schulz & Grant, 2000). Facilitating transfer of skills throughout settings is another crucial aspect of therapy to ensure meaningful results for the client.

Other intervention methods used during speech therapy for PD include voice amplifiers, delayed auditory feedback (DAF) devices, intensity biofeedback devices, masking devices, and expiratory muscle strength training (EMST) (Schulz & Grant, 2000; Baker et. al, 2005). Voice amplification increases the vocal loudness of PD speech without requiring much effort from the speaker. DAF and biofeedback devices both provide the speaker with speech information regarding intelligibility, loudness, rate, and pitch. Speakers who wear DAF or biofeedback devices are taught during therapy how to make adjustments in speech production related to the speech information they receive from the device. EMST increases the output force of the expiratory muscles and allows for improvements in voice, coughing, and swallowing (Baker et. al, 2005).

#### Acoustic Applications for PD

Not only are the professionals in the field of communicative sciences and disorders looking to improve the communication of those with PD, but they are also looking for ways to more precisely define PD speech. The increasing wealth of literature providing knowledge of the detailed properties of PD speech offers opportunities to investigate the use of these variables for early detection, differential diagnosis, and monitoring. It is practical to explore the use of PD

speech for potential disease detection and monitoring because it is inexpensive, noninvasive, and can be used through means of telepractice.

Traditionally, perceptual measures have been used to describe the characteristics of hypokinetic dysarthria that accompany PD. Unfortunately, perceptual measures are at risk for observer bias and do not typically allow for precision in measuring change. In the late 1990's, Kent and colleagues (1999) recognized the need for objective, acoustic applications for the assessment of dysarthria. Early detection and differential diagnosis using acoustic applications would allow for earlier treatment of PD, which may slow disease progression and increase quality of life. Now, progress in technology has made this goal feasible through use of assorted techniques.

Digitizing Sound for Acoustic Applications. There are copious equipment and software options available for recording sound to use in acoustic analysis. Regardless of the type of microphone or computer software used, sound is digitized through the same general process. Sound originating from the talker in the form of mechanical energy is received by an acoustical receiver, usually a microphone, and is converted into electrical energy through a transducer. During the conversion, the sound waveform is sampled and converted into a string of numbers so it can be stored digitally on the computer.

WAV files are one format for storing audio files digitally and are often used for acoustic analysis applications. WAV files are uncompressed, discrete time series which represent the pressure wave of the sound, or sounds, of interest. Elements

used to recreate the sound in WAV files are sampling rate, Nyquist frequency, amplitude resolution, channels, and type of compression.

Sampling rate is measured in cycles or samples per second and reproduces the continuous time signal. Universally, the Nyquist frequency is half of the sampling rate. Aliasing, which is the distortion of the recreated audio signal, occurs when frequencies are greater than the Nyquist frequency. To avoid this, the maximum frequency of the continuous-time signal of interest should be less than the Nyquist frequency.

The amplitude resolution of WAV files is typically either 8 bits or 16 bits. Eight bits are considered to be one byte and 16 bits are considered to be two bytes. The size of a WAV file can be calculated by multiplying the duration of the recording in seconds by the sampling rate, then multiplying the product by the amplitude resolution and converting the units into bytes. The digitized signal can then be stored and recreated as needed for analysis but investigators should be knowledgeable about the recreated digital signal. Quantization noise is a phenomenon that occurs when information is added to the recreated sound since resolution isn't high enough to exactly recreate the signal. Essentially, quantization noise limits the precision to which a digital signal can represent the original signal from the talker (Denes & Pinson, 1993). Quantization noise is present in all digital signals, but 16 bits of resolution has less quantization noise than 8 bits of resolution.

Channels in a WAV file refer to the number of microphones used to record speech. Two channel WAV files will have had two microphones to complete the recording. Single channel will have had one microphone. The number of channels used to record is determined by the needs of the acoustic application being conducted.

WAV files are compressed in various ways to allow for optimal storage of digital files. There are two types of compression: amplitude compression and data compression. Amplitude compression is performed to avoid clipping. In data compression, files can be lossy or lossless. Lossy compression is achieved by getting rid of information in the signal that is not perceived to be as important as other parts of the signal. MP3 files are lossy compressed to allow for maximum storage and listening pleasure. Lossless compression is the standard for speech analysis because files are stored in a format which returns to the exact same information as before compression.

Spectral-Based Analysis. Once speech recordings are stored digitally, temporal or spectral-based analyses can be conducted. To achieve this, it is necessary to select a portion of the waveform to use for temporal and amplitude measurements. The intervals selected for analysis are called frames (Kent & Read, 2002). Frame length is the term for the duration of the frame, and is typically 20-30 ms (Kent & Read, 2002). Frame-based analysis is a method for analyzing speech in a continuous, overlapping way so that one can observe speech in moments and over time. The extent to which frames overlap is called the frame interval. Frame

intervals require precise calculation as the analysis may miss important information in the signal if overlap is too small, and may perform unnecessary computation if overlap is too large (Kent & Read, 2002).

Windows are applied to the signal during frame-based analysis. Kent and Read (2002) report that a window is "...a weighting function applied to a waveform so that its amplitude is shaped in a particular fashion...to minimize the amplitude at the edges of the window so that the analysis focuses on a representative part of the signal". Rectangular windows are commonly used due to their simplicity, but there are a couple trade-offs to using this window; the temporal characteristics are preserved but the waveform is truncated at the boundaries (Deller, Jr., et. al, 2000). Hamming, Hanning, and Blackman Windows allow for smoother truncations, though they slightly distort the temporal waveform (Deller, Jr., et. al, 2000). Experts report that smoother windows are typically chosen for analysis due to their preferable side-lobe characteristics.

Following frame and window selection, the waveform in the speech recording can be adapted into various displays so that the analyses of interest can be completed. Discrete Fourier Transform (DFT) is a mathematical process that translates the time domain values of a waveform in digital data into frequency domain values of a spectrum (Kent & Read, 2002). Fast Fourier Transform (FFT) is a special type of DFT which is commonly used on computers to create a Fourier spectrum (Kent & Read, 2002). FFT is the basis for many common speech analysis methods, including formant estimators, pitch estimators, spectral measures, and

cepstral measures. Many computer software programs are available that are capable of preparing digital sound for spectral analysis and investigating variables of interest, including those related to PD.

#### **Recording Speech for Analysis**

A full evaluation of speech function is best completed by administering several vocal tests. Vocal tests are chosen based on study design and variables of interest. Recording speech for use in research and clinical applications is favorable due to relative ease for participant to complete, potential use in telepractice, and ability save records and use for multiple purposes.

The recording of conversational speech will reflect data for speech's natural purpose and may be used for measures such as rate, prosody, and coarticulation. Researchers should use caution when employing conversational speech tasks because they run a high risk of Hawthorne effects which may confound the results. Hawthorne effects occur when a study participant consciously changes the way he is speaking when he knows he is being recorded (Lansberger, 1959). Still, the recording of conversational speech can be useful if procedures are established to control for confounding factors.

A vocal task similar to conversational speech in utterance length which may control for Hawthorne effects and differences between samples is running speech. In running speech tasks, study participants read a sentence or sentence group that has been constructed to cater to the interests of the specific study (Little et al., 2009). Running speech tasks offer a way to control and ensure uniformity in data,

but they are also vulnerable to confounding factors as well, such as the linguistic components of speech (Little et al., 2009).

Vocal tasks that are not representative of speech in language use can also be beneficial to researchers depending on what measure is of interest. Sustained phonations of vowels and diadochokinetic tasks offer methods of further eliciting characteristics of speech (Little et al., 2009). Again, these tasks are easy to perform and offer the opportunity to investigate the many characteristics of speech.

#### Acoustic Measures for PD

Voice Quality Measures. The American National Standards Institute (ANSI) states that sound quality is "...that attribute of auditory sensation in terms of which a listener can judge that two sounds, similarly presented and having the same loudness and pitch, are dissimilar" (Titze & Verdolini Abbott, 2012). Perceptual measures of voice quality are usually completed by recording judgments of vocal roughness, breathiness, strain, pitch, and loudness. PD voice quality rating tasks are typically interested in breathiness, pitch, and loudness. These ratings are typically categorized as normal, mild, moderate, and severe.

Though perceptual measures of voice can provide meaningful information to professionals of interest, there are many factors that can affect the precision of these measurements. Perceptual measures of voice can be inconsistent across listeners. The vocal perception of a listener is influenced by his individual internal references for voice quality attributes which can change over time and be prejudiced by voices just heard (Titze & Verdolini Abbott, 2012). The use of perceptual voice

quality measurements for PD speech and other pathologies still have a place in the realms of clinics and laboratories today, but several acoustic measures of voice quality have been developed since the recommendation of acoustic methods for measuring speech by Kent and colleagues in 1999.

Pitch Period Entropy (PPE) is a robust measure of dysphonia which has been found to be accurate in classifying PD voices from healthy voices. In their 2009 study, Little and colleagues found that PPE was 91.4% accurate in classifying the sustained phonations of people with PD from those of controls. This variable is not only important because of its sensitivity to PD, but also because it is robust to many factors in the acoustic environment that are difficult to control for (Little et al., 2009). The characteristics of PPE make it an important tool for the clinic and telepractice.

Fundamental frequency (F0) and its derivatives are also useful measures of voice quality of PD speech. F0 is the lowest frequency in a periodic wave and its use allows for quantitative measurement of phonation and vocal quality (Denes & Pinson, 1993). There are several F0 statistics available for voice analysis, including F0 mean, mode, range, and standard deviation (Kent et al., 1999). The standard deviation of F0 (oF0) is a long-term measure of phonatory instability. Kent and colleagues (1999) report that oF0 is more useful than measures of shimmer and jitter, perturbation measures, for distinguishing speakers with PD from typical speakers. oF0 is also used to measure the linear decline of voice in speakers with PD.

The Auditory-Based Spectrum-Enhancing Preprocessing Stage Sawtooth Waveform Inspired Pitch Estimator (AUD-SWIPE) is a relatively new system for quantifying pitch characteristics in speech. In brief, AUD-SWIPE returns pitch and pitch strength values by passing the signal through an auditory-based spectrum enhancing pre-processing stage, which is then fed into the pitch estimator (Camacho, 2012). AUD-SWIPE is capable of calculating the mean and standard deviation (std), among other values, for pitch (PT) in Hertz (Hz) and semitones (ST), as well as pitch strength (PS). PT is the subjective quality that describes the frequency of a signal, such as F0; the higher the frequency the higher the perceived PT (Denes & Pinson, 1993).

Acoustic measures of PT and PS are more advantageous than F0 measures and PPE because they are more representative of the speech signal as a whole and can be applied to all types of speech recordings, not just sustained vowels. PT is preferred over F0 because PT represents a perception which is found to be, in general, more relevant in acoustic research than F0 which represents a physical acoustic property. Further, measuring PT in ST is advantageous to Hz because ST is more representative of the normal distribution of PT. Figure 1 shows the transformation from Hz to ST. The transformation is a base 2 log function with reference frequency C0 = 16.352 Hz (C4 is "middle C" in the equal temperament musical tuning, four octaves above C0, and A4 = 440 Hz is A above middle C). The scale is such that 12 semitones = 1 octave.



Figure 1: Frequency in Semitones compared to Frequency in Hz

Cepstral peak prominence (CPP) is a measure of dysphonia that represents F0 and can be applied to continuous voice samples (Watts & Awan, 2011). CPP is measured by taking the difference in amplitude between the cepstral peak and the corresponding value in the regression line that is directly below the peak (Hillenbrand et al., 1994). It represents how far the cepstral peak emerges from the cepstral background noise (Hillenbrand et al., 1994). CPP is sensitive and specific for dysphonia, and it is highly related to perceptions of voice quality in hypofunctional speakers. Whereas the validity of other measures, such as jitter and shimmer, becomes questionable as the dysphonia worsens, CPP is capable of being valid when applied to severely dysphonic voices (Watts & Awan, 2011).

The significance of CPP was demonstrated in a 2011 study conducted by Watts and Awan. Researchers used the measure to differentiate dysphonic voices from healthy voices using sustained vowels and continuous speech samples. Results demonstrated that CPP was significantly different between groups in both speaking tasks (Watts & Awan, 2011). Further analyses of dysphonia demonstrated high sensitivity and high specificity for the CPP in the sustained vowel task and high sensitivity and moderate specificity for the CPP in the continuous speech task (Watts & Awan, 2011), thus making it an important measure for the detection and management of PD.

Articulation Measures. Perceptual measures of articulation also have purpose in voice clinics and laboratories. Unlike voice quality measures that can be judged by almost any listener, perceptual measures of articulation probably require the listener to have knowledge of the International Phonetic Alphabet and phonetic transcription. Use of phonetic transcription and diacritical markers allows for the perceptual judgment of articulation, but it is subject to listener bias and inconsistency across raters just as any perceptual measure would be. Terms such as slurred, slushy, and imprecise have been used to describe listeners' perception of PD speech, but they do not offer detailed information as phonetic transcription would. Acoustic measurements of articulation offer a means of evaluating the production of vowels and consonants so that the issues in perceptual measures are controlled for.

Since vowels and consonants are produced by manipulation of structures in the vocal track, the restriction of movement seen in PD is reflected by articulation in speech production. Some recent, innovative measurements of articulation are

human factor cepstral coefficients (HFCCs) and their variations (Skowronski & Harris, 2004). HFCCs, acoustic correlates of articulations, are derived from mel frequency cepstral coefficients (MFCCs) which are measured from a sound cepstrum. MFCCs have been commonly used in many speech research domains, including automatic speech recognition, speech synthesis, and speech coding, because they represent the perceptually relevant aspects of the vocal tract and offer robustness to noise in recordings (Davis & Mermelstein, 1980). A study conducted by Blocket and colleagues (2011) found MFCCs to be 88% accurate in recognizing PD speech.

HFCCs have been demonstrated to have improved noise robustness in automatic speech recognition experiments over MFCCs (Skowronski & Harris, 2004). Skowronski and colleagues (2012) report that cepstral coefficient standard deviation sum (CC SDS) and delta cepstral coefficient standard deviation sum (DCC SDS) demonstrated 92.9% accuracy in discriminating PD speech from speech of healthy talkers. The following equations are used to determine the cepstral coefficient measures:

$$CC SDS = \sum_{k=1}^{K} \sigma_{CC}(k)$$

$$\Delta \text{CC SDS} = \sum_{k=1}^{K} \sigma_{\Delta \text{CC}}(k)$$

where ( $\Sigma$ ) is the sum, (k) is the cepstral index, (K) is the total number of cepstral coefficients, and (o) is the standard deviation of CC or  $\Delta$ CC.

Outcomes of the 2012 study revealed several findings. First, the duration of sentence utterance, duration of active speech, and duration of pauses were all significantly longer for PD speakers compared to healthy speakers. Next, the activity factor was significantly lower for PD speakers (Skowronski et al., 2012). Authors propose that the combination of HFCC variations with voice and prosodic variables may further identify PD speech and may help distinguish it among the other dysarthrias.

#### The Present Study

The purpose of this study was to investigate the usefulness of voice and articulation measures for distinguishing PD speech before clinical diagnosis and demonstrating PD progression over time. The research questions addressed by this study were:

- 1. Do the variables of interest successfully differentiate PD speakers from healthy speakers at time points before a clinical diagnosis of PD?
- 2. Can the variables of interest be used to demonstrate PD progression over several decades?

To answer the stated questions, a study was designed to employ the aforementioned variables to compare the voice and articulation functions of speakers with PD, before and after diagnosis, to a population of age- and gender-matched controls. The specific voice and articulation variables used in this study were PT mean ST, PT std ST, PS mean, CPP, CC SDS, and DCC SDS.

First, it is hypothesized the mean values of the dependent variables will be significantly different between the three speaker groups; with the values for the PD Post-Dx group being the lowest and the values for the Control group being the greatest. The first hypothesis is based on knowledge that PD speech is characterized by mono-pitch, which would yield a lower PT std, hoarse or harsh voice, which would yield lower CPP and PS mean measures, as well as hypokinetic dysarthria, which would yield lower CC SDS and DCC SDS measures. Second, it is hypothesized that all dependent variables will decline over time due to effects of aging, and that the variables for PD speakers will decline more rapidly due to disease effects; with the PD Post-Dx group having the most severe decline and the Control group having the least severe decline.

#### METHODOLOGY

#### **Design** Type

A retrospective, longitudinal design was used for this study.

#### Participants

Ten speakers with a clinical diagnosis of Parkinson's disease and nine healthy control typical speakers were used for this study. Speakers were included in the study if speech recordings from multiple time points were available for analysis. Three participant groups were used for comparison: PD speakers before clinical diagnosis (PD Pre-Dx), PD speakers after clinical diagnosis (PD Post-Dx), and healthy speakers (Control).

#### **Data Collection**

Three PD participants were recruited from the university clinic and surrounding community by means of flyer advertisements and promotion of the study at local PD support groups. Participants were included in the study if they were at least 50 years old, had a clinical diagnosis of PD, and could provide recordings of their speech from multiple time points before their diagnosis.

Participants from the community attended one recording session at the Voice Acoustics and Perception Laboratory (VAPL) at Michigan State University where they agreed to the informed consent and authorized a medical release form in accordance with the institutional review board at Michigan State University. Before performing speech tasks, all participants were screened for hearing ability bilaterally at 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz (thresholds below 25 dB HL).

Recordings were conducted in a sound-proof booth using a TASCAM DR-40 Linear PCM Recorder which was held in front of the participant by the researcher. Recordings of conversational speech were obtained by asking the participant interview-like questions. Participants provided the researcher with personal media files to be used for analysis at time of the lab recording session. Participants were compensated \$20 for their participation.

Media files provided by participants recruited from the community were in the form of VHS tapes, DVDs of converted VHS tapes, and one Olympus VN-480 PC Digital Voice Recorder. WAV files were extracted from the VHS tapes by means of the computer's sound card. A Samsung TV, model number CXD1942, with a builtin VCR was connected to the computer's sound card. Next, Audacity was used to record the signal transmitted through the sound card and convert it into singlechannel WAV files with sampling rates of 44.1 kHz and resolutions of 32 bits. WAV files were extracted from the DVDs by means of ZC DVD Audio Ripper software, a freeware available online. Software specific to the Olympus VN-480 PC Digital Voice Recorder was provided for download from an Olympus product support technician. WAV files were then copied and pasted from the recorder to laboratory computers. All WAV files obtained from all three participants' personal media were later segmented for the target speaker using PRAAT.

Internet searches for public figures with a known diagnosis yielded seven participants with PD. Recordings for six of these participants were obtained using the Vincent Voice Library (VVL) at Michigan State University. Requests were

submitted for recordings made at multiple time points for each speaker. Recordings were segmented for target speakers using PRAAT. For all six participants, postdiagnosis recordings were obtained using the VVL and YouTube.

Recordings for the tenth PD participant were obtained using the Illinois Supreme Court website's Supreme Court Oral Argument Audio & Video database. A YouTube clip provided a post-diagnosis recording for this participant. Recordings were segmented for the target speaker using PRAAT.

Age- and gender-matched controls were found by completing internet searches for politicians and reporters born in the same years as the PD speakers they were to be matched to. Names found through the internet search were then entered into the VVL search bar to see if any recordings were available. A second search was conducted on YouTube if the VVL search did not yield enough recordings. Age- and gender-matched controls were included in the study if the searches yielded recordings for the control speaker near the same time points as the PD speaker. This method yielded nine age- and gender-matched controls

Varieties and time ranges of recordings are summarized in Tables 1, 2 and 3, and in Figures 2 and 3.

	Lab Recordings	Personal Audio Recordings	VVL Recordings	YouTube Recordings	Other Internet Recordings	Total
PD Pre-Dx		12	19	1	3	38
PD Post-Dx	3	3	2	10		15
Control			43	9		52
Total	3	15	64	20	3	105

Table 1: Sources Used to Obtain Recordings

	Conversational /Casual	Formal Interview	Speech	Report	Lab Recording	Total	Range of Recordings in Years
PD1	5				1	6	25
PD2	4	1			1	6	36
PD3	3	1				4	5
PD4		2	7			9	59
PD5	5				1	6	27
PD6		1	1			2	8
PD7		3	2			5	8
PD8		1	4			5	14
PD9		4	1			5	10
PD10		4	1			5	30
Total	17	17	16		3	53	61

Table 2: Types of Recordings Used in Analysis and Time Range for PD Speakers

	Conversational /Casual	Formal Interview	Speech	Report	Lab Recording	Total	Range of Recordings in Years
C1		1	3			4	22
C2		4	2			6	38
C3			<b>5</b>			<b>5</b>	5
C4		2		5		7	46
C5		1	<b>5</b>			6	25
C6		4	1	1		6	41
<b>C</b> 7		1	2	<b>5</b>		8	55
C8		4	1			5	20
<b>C9</b>		4	1			5	24
Total		21	20	11		52	74

Table 3: Types of Recordings Used in Analysis and Time Range for Control Speakers



# Range of Recording Time Points- PD Participants

Figure 2: Time Range of Recordings and Time Points of Recordings Used for Analysis- PD Participants



Range of Recording Time Points- Control Participa Chart Area

Figure 3: Time Range of Recordings and Time Points of Recordings Used for Analysis- Control Participants

#### Procedure

Data was analyzed using MATLAB software version 8.2.0.701 (R2013b). Scripts in MATLAB used for data analysis were a hand-label GUI with routines for HFCC measures and AUD-SWIPE. Files with sampling rates greater than 44.1 kHz were re-sampled using a GUI feature in MATLAB. WAV files of PD participants and controls were cut into 5-second "snippets" using the automatic function available in the GUI. PT mean ST, PT std ST, and PS mean values were obtained using AUD-SWIPE. Data for female talkers was removed for the PT mean ST measure so they would not act as outliers in the output values. CC SDS and DCC SDS measures were made using the hand-label HFCC GUI. CPP was measured using code created by Hillenbrand (Hillenbrand et. al, 1994). Variable values yielded from the 5-second snippets were averaged per recording time point per speaker.

1-way ANCOVA was conducted for each dependent variable for all three speaker groups to determine the regression line slope for the Control model, the residual measures of the PD Pre-Dx and Post-Dx groups, and the regression line slopes for the PD Pre-Dx and Post-Dx residual measures. Post hoc testing was conducted to determine if there were significant differences among slope values.

1-way ANOVA was conducted for each dependent variable for all three speaker groups: PD Pre-Dx, PD Post-Dx, and Control. Post hoc testing was conducted to determine if there were statistical differences among group means.

The data for the PD Pre-Dx and PD Post-Dx groups were combined to increase statistical power and a second 1-way ANCOVA with post hoc testing and 1way ANOVA with post hoc testing was performed for each variable. Only measures with significant results for the combined PD talker group (PD Combined) are reported in the results section.

#### RESULTS

PT mean ST: 1-way ANCOVA with Post Hoc Testing. 1-way ANCOVA with post hoc testing revealed that the regression line slope (RLS) for the PD Pre-Dx group residual measure was significantly less than the Control group RLS (p<0.05). The RLS for the PD Post-Dx group was not significantly different from the RLS of the other talker groups. Figure 4 depicts the regression lines for all three speaker groups, with the black line representing the Control group, the blue line representing the PD Pre-Dx group, and the red line representing the PD Post-Dx group. The Control RLS has been adjusted to fit a slope of 0 for all measures to account for the influence of age, thus allowing a better representation of RLS affected by PD.

Table 4 represents the 1-way ANCOVA test results where d.f. represents the degrees of freedom, F is the f-statistic, and Prob>F is the probability that there is a significant difference in RLS among talker groups.

Figure 5 represents post hoc testing for the 1-way ANCOVA of PT mean ST, where the red circle depicts the RLS of the data and the interval bars, derived from Tukey Kramar honest significant difference, are the range of which the RLSs exist.



Figure 4: Regression Line Slopes- PT mean ST

Source	d.f.	F	Prob>F
Speaker Group	2	3.58	p<0.05
Age	1	2.21	p=0.14
Speaker Group * Age	2	3.17	p<0.05
Error	90		

Table 4: ANCOVA Test Results- PT mean ST



PT mean ST: 1-way ANOVA with Post Hoc Testing. 1-way ANOVA with post hoc testing revealed that the PT mean ST average of the PD Pre-Dx group was significantly greater than that of the Control group (p<0.05). The PT mean ST average for the PD Post-Dx group was not significantly different from any other talker group. Figure 6 represents post hoc testing for 1-way ANOVA of PT mean<sup>‡</sup> ST, where the red circle depicts the sample mean of the data and the interval bars, derived from Tukey Kramar honest significant difference, are the range of which the true mean exists. 1-way ANOVA test results for PT mean ST are seen in Table 5, where d.f. stands for degrees of freedom, F represents the f-statistic, and Prob>F is the probability there is a significant difference among group means.



Figure 6: ANOVA Post Hoc Test- PT mean ST

Source	d.f.	F	Prob>F
Speaker Group	2	3.36	p<0.05
Error	93		
Total	95		

Table 5: ANOVA Test Results- PT mean ST

*PT std ST: 1-way ANCOVA with Post Hoc Testing.* 1-way ANCOVA with post hoc testing revealed that there were no significant differences in PT std ST RLS among the three talker groups, as indicated by the ANCOVA test results seen in Table 6 and the post hoc test depicted in Figure 8. Figure 7 illustrates the RLS for all talker groups.



Figure 7: Regression Line Slopes- PT std ST

Source	d.f.	F	Prob>F
Speaker Group	2	0.49	p=0.62
Age	1	0.03	p=0.87
Speaker Group * Age	2	0.9	p=0.41
Error	100		

Table 6: ANCOVA Test Results- PT std ST



PT std ST: 1-way ANOVA with Post Hoc Testing. 1-way ANOVA with post

hoc testing revealed that there were no significant differences among PT std ST group means for the three talker groups, as indicated by the ANOVA post hoc test depicted in Figure 9 and the ANOVA test results seen in Table 7.



Figure 9: ANOVA Post Hoc Test- PT std ST

Source	d.f.	F	Prob>F
Speaker Group	2	0.48	p=0.62
Error	103		
Total	105		

Table 7: ANOVA Test Results- PT std ST

*PS mean: 1-way ANCOVA with Post Hoc Testing.* 1-way ANCOVA with post hoc testing revealed that there were no significant differences in PS mean RLS among the three talker groups, as indicated by the ANCOVA test results seen in Table 8 and the post hoc test depicted in Figure 11. Figure 10 illustrates the RLS for all talker groups.



Figure 10: Regression Line Slopes- PS mean

Source	d.f.	F	Prob>F
Speaker Group	2	0.71	p=0.49
Age	1	0.38	p=0.54
Speaker Group * Age	2	1.43	p=0.24
Error	100		

Table 8: ANCOVA Test Results- PS mean



PS mean: 1-way ANOVA with Post Hoc Testing. 1-way ANOVA with post

hoc testing revealed that there were no significant differences among PS mean group means for all three talker groups, as indicated by the ANOVA post hoc test in Figure 12 and the ANOVA test results seen in Table 9.



Figure 12: ANOVA Post Hoc Test- PS mean

Source	d.f.	F	Prob>F
Speaker Group	2	0.69	p=0.51
Error	103		
Total	105		

Table 9: ANOVA Test Results- PS mean

*CPP: 1-way ANCOVA with Post Hoc Testing.* 1-way ANCOVA with post hoc testing revealed that the CPP RLS for the PD Pre-Dx group was significantly less than the RLS for the Control group (p<0.05), as indicated by the ANCOVA test results seen in Table 10 and the post hoc test depicted in Figure 14. The RLS for the PD Post-Dx group was not significantly different from the RLS of the other talker groups. Figure 13 illustrated the RLS for all talker groups.



Figure 13: Regression Line Slopes- CPP

Source	d.f.	F	Prob>F
Speaker Group	2	2.61	p<0.05
Age	1	2.71	p=0.10
Speaker Group * Age	2	2.92	p=0.058
Error	100		

Table 10: ANCOVA Test Results- CPP



CPP: 1-way ANOVA with Post Hoc Testing. 1-way ANOVA with post hoc testing revealed that the CPP group means for the Control and PD Pre-Dx groups were significantly greater than the CPP group mean for the PD Post-Dx group (p<0.05), as indicated by the ANOVA post hoc test in Figure 15 and the ANOVA test results seen in Table 11.



Source	d.f.	F	Prob>F
Speaker Group	2	4.46	p<0.05
Error	103		
Total	105		

Table 11: ANOVA Test Results- CPP

CC SDS: 1-way ANCOVA with Post Hoc Testing. 1-way ANCOVA with post hoc testing revealed that there were no significant differences among the CC SDS RLS for the three talker groups, as indicated by post hoc testing seen in Figure 17. The CC SDS RLS for all three talker groups are illustrated in Figure 16 and the ANCOVA test results can be seen in Table 12.



Figure 16: Regression Line Slopes- CC SDS

Source	d.f.	F	Prob>F
Speaker Group	2	5.32	p<0.05
Age	1	0.42	p=0.52
Speaker Group * Age	2	0.52	p=0.59
Error	100		

Table 12: ANCOVA Test Results- CC SDS



*CC SDS: 1-way ANOVA with Post Hoc Testing.* 1-way ANOVA with post hoc testing revealed that the CC SDS group mean for the PD Pre-Dx group was significantly less than the CC SDS group mean for the Control group (p<0.05), as indicated by post hoc testing depicted in Figure 18 and ANOVA test results seen in Table 13.



Source	d.f.	F	Prob>F
Speaker Group	2	6.43	p<0.05
Error	103		
Total	105		

Table 13: ANOVA Test Results- CC SDS

DCC SDS: 1-way ANCOVA with Post Hoc Testing. 1-way ANCOVA with post hoc testing revealed that there was a significant difference among speaker groups (p<0.05) but not among RLS, as indicated by ANCOVA test results seen in Table 14 and post hoc testing depicted in Figure 20. The DCC SDS RLS for all three talker groups are illustrated in Figure 19.



Figure 19: Regression Line Slopes- DCC SDS

Source	d.f.	F	Prob>F
Speaker Group	2	5.71	p<0.05
Age	1	0.01	p=0.93
Speaker Group * Age	2	0.59	p=0.56
Error	100		

Table 14: ANCOVA Test Results- DCC SDS



*DCC SDS: 1-way ANOVA with Post Hoc Testing.* 1-way ANOVA with post hoc testing revealed that the DCC SDS group mean for the PD Pre-Dx group was significantly less than the DCC SDS group mean for the Control group (p<0.05), as indicated by post hoc testing depicted in Figure 21 and ANOVA test results seen in Table 15.



Figure 21: ANOVA Post Hoc Test- DCC SDS

Source	d.f.	F	Prob>F
Speaker Group	2	0.99	p<0.05
Error	103	0.15	
Total	105		

Table 15: ANOVA Test Results- DCC SDS

Combined PD Talker Group: 1-way ANCOVA with Post Hoc Testing. 1-way ANCOVA with post hoc testing revealed that the CPP RLS for the PD group was significantly less than the CPP RLS for the Control group (p<0.05), as indicated by post hoc testing seen in Figure 23 and ANCOVA test results seen in Table 16. The CPP RLS for the two talker groups are illustrated in Figure 22.



Figure 22: Combined PD Group Regression Line Slopes- CPP

Source	d.f.	F	Prob>F
Speaker Group	1	2.55	p=0.11
Age	1	7.26	p<0.05
Speaker Group * Age	1	8.53	p<0.05
Error	102		

Table 16: ANCOVA Test Results- CPP



Combined PD Talker Group: 1-way ANOVA with Post Hoc Testing. 1-way ANOVA with post hoc testing revealed that the PT mean ST group mean for the PD group was significantly greater than the PT mean ST group mean for the Control group (p<0.05), as indicated by post hoc testing depicted in Figure 24 and ANOVA test results seen in Table 17.

1-way ANOVA with post hoc testing revealed that the CC SDS group mean for the PD group was significantly less than the CC SDS group mean for the Control group (p<0.05), as indicated by post hoc testing depicted in Figure 25 and ANOVA test results seen in Table 18.

1-way ANOVA with post hoc testing revealed that the DCC SDS group mean for the PD group was significantly less than the DCC SDS group mean for the Control group (p<0.05), as indicted by post hoc testing depicted in Figure 26 and ANOVA test results seen in Table 19.



Figure 24: Combined PD Group ANOVA Post Hoc Test- PT mean ST

Source	d.f.	F	Prob>F
Speaker Group	1	6.71	p<0.05
Error	94		
Total	95		

Table 17: ANOVA Test Results- PT mean ST



Figure 25: Combined PD Group ANOVA Post Hoc Test- CC SDS

Source	d.f.	F	Prob>F
Speaker Group	1	8.53	p<0.05
Error	104		
Total	105		

Table 18: ANOVA Test Results- CC SDS



Figure 26: Combined PD Group ANOVA Post Hoc Test- DCC SDS

Source	d.f.	F	Prob>F
Speaker Group	1	5.23	p<0.05
Error	104		
Total	105		

Table 19: ANOVA Test Results- DCC SDS

#### DISCUSSION

The intention of the present study was to investigate the use of acoustic variables (PT mean ST, PT std ST, PS mean, CPP, CC SDS, DCC SDS) for distinguishing PD speakers from healthy speakers before a clinical diagnosis of PD and for demonstrating disease progression over several decades. The first hypothesis was that the mean values of the acoustic variables would be significantly different between the three speaker groups (PD Pre-Dx, PD Post-Dx, and Control), with the PD Post-Dx group having the smallest values and the Control group having the greatest values. The second hypothesis was that the acoustic variables would decline more rapidly over time for PD speakers due to aging effects exacerbated by PD, with the PD Post-Dx group having the most severe decline and the Control group having the least severe decline.

Differences in Group Means Determined by ANOVA. No variables supported the first hypothesis. However, significant differences among group mean values were found for four variables following ANOVA. The PT mean ST group mean for the PD Pre-Dx group was significantly greater than the PT mean ST group mean for the Control group. The CPP group means for the Control and PD Pre-Dx groups were significantly greater than the CPP group mean for the PD Post-Dx group. The CC SDS group mean for the PD Pre-Dx group was significantly less than the CC SDS group mean for the Control group. The DCC SDS group mean for the PD Pre-Dx group was significantly less than the DCC SDS group mean for the Control group. No significant differences among group mean values were found for the PT std ST and PS mean measures.

Following combination of the PD Pre-Dx and PD Post-Dx groups into the PD group, ANOVA analysis revealed significant differences between group mean values for three variables. The PT mean ST group mean for the PD group was significantly greater than the PT mean ST group mean for the Control group. The CC SDS and DCC SDS group means for the PD group were significantly less than the CC SDS and DCC SDS group means for the Control group.

Differences in Regression Lines Determined by ANCOVA. No measures supported the second hypothesis regarding variable decline severity in the three speaker groups. Significant differences among RLS for talker groups were found for two variables. The PT mean ST RLS for the PD Pre-Dx group is significantly less than the PT mean ST RLS for the Control group. The CPP RLS for the PD Pre-Dx group was significantly less than the CPP RLS for the Control group. No significant differences were found among group RLS for the PT std ST, PS mean, CC SDS, and DCC SDS measures.

Following combination of the PD Pre-Dx and PD Post-Dx groups into the PD group, ANCOVA analysis revealed significant differences between group RLS for one variable. The CPP RLS for the PD group was significantly less than the CPP RLS for the Control group.

*Implications of Study Findings.* Though the original hypotheses of this study were not supported by any of the six measures, the results reveal some remarkable

findings. First, the PT mean ST mean was found to be significantly greater in the PD Pre-Dx group and combined PD group than the Control group. Preliminary research conducted by Hunter and colleagues suggests that over one's life time, PT will decrease with age but then begin to increase around the mid to latter part of the sixth decade (Hunter & Banks, 2014). This preliminary data could be what's influencing the outcomes of the PT mean ST data in the present study.

Second, the CC SDS and DCC SDS group mean values of the PD Pre-Dx and the combined PD groups were significantly less than those of the Control group. This demonstrates that HFCC measures can be used to differentiate PD talkers from healthy talkers at times before and after clinical diagnosis.

Finally, the CPP RLS for the PD Pre-Dx group and the combined PD group were significantly less than the RLS for the Control group. This demonstrates that CPP can be used as an indicator of vocal quality decline as a result of PD.

One question raised by the outcomes of this study is in regards to the PD Post-Dx group. Why wasn't the group mean or RLS of the PD Post-Dx group significantly less than those of the Control group for any measure? It is difficult to interpret the meaning of clinical diagnosis, as this is an arbitrary title for when the talker knew PD was present, but the term does not provide information as to the onset of the disease. The measures for which the PD Pre-Dx group values, but not the PD Post-Dx values, were significantly less than the Control group values could be explained by initiation and implementation of treatment methods following diagnosis, speaker compensation due to awareness of PD, or increased quality of

recording technology as is developed over time, all of which may account for increased vocal performance in the PD Post-Dx data.

*Limitations.* Though some aspects of the results can accounted for by previously stated explanations, other aspects cannot be clarified. The unexpected outcomes of the present study are likely confounded by its retrospective design-type, as retrospective research can have issues with data reliability and validity. It is impossible to tell certain aspects of the recordings used in analysis, such as digitizing settings, environment, and quality of equipment used for recording. Also, the small sample size may not provide a representative sample of the population, which is why data outcomes were not what was anticipated. The poor control of recordings and limited pool of participants to obtain recordings from have likely influenced the study's outcomes.

Suggestions for Future Research. Outcomes of longitudinal studies of PD speech may be improved should this study be recreated in such a way that recordings are controlled for digitizing settings and recording environments. Also, precise, longitudinal data collection of individuals at risk for PD may provide insight to PD onset, which benefit research regarding early disease detection and progression monitoring.

APPENDIX

### APPENDIX

Speaker	Recording Type and Year
PD1	C/C-1988, 1989, 1990, 1991, 1992; Lab-2013
PD2	C/C-1977, 1978, 1979, 1980; FI-1996; Lab-2013
PD3	C/C-2008, 2009, 2010; Lab-2013
PD4	Speech-1952, 1963, 1976, 1986, 1990, 1998, 2001, 2006; FI-2011
PD5	C/C- 1986, 1992, 2006, 2008, 2009; Lab-2013
PD6	Speech-1968; FI-1976
PD7	Speech- 1969, 1975; FI- 1974, 1976, 1977
PD8	Speech- 1993, 1997, 2000, 2003; FI- 2007
PD9	FI- 1974, 1975, 1978, 1984; Speech- 1980
PD10	Speech- 1966; FI- 1974, 1975, 1977, 1996
C1	FI- 1984; Speech- 1991, 2004, 2006
C2	FI- 1970, 1975, 1984, 2005; Speech- 1996, 2008
C3	Speech <sup>-</sup> 2008, 2009, 2010, 2012, 2013
C4	Report- 1954, 1960, 1962, 1977, 1981; FI- 1996, 2000
C5	Speech- 1988, 1996, 2008, 2012, 2013; FI- 2004
C6	FI- 1961, 1974, 1983, 2002; Speech- 1976; Report- 1997
C7	Report- 1939, 1958, 1963, 1968; Speech- 1973, 1978, 1983, 1994
C8	Speech- 1993; FI- 1997, 2008, 2010, 2013
C9	FI- 1975, 1976, 1978,1999; Speech- 1997

Table 20: Type of Recording and Year for Individual Speakers

Key: C/C= Conversational/Casual, FI=Formal Interview, Lab= Lab Recording,

Speech= Formal Speech, Report= Report performed in Studio

REFERENCES

#### REFERENCES

- Baker, S., Davenport, P., & Sapienza, C. (2005). Examination of Strength Training and Detraining Effects in Expiratory Muscles. *Journal of Speech, Language,* and Hearing Research, 48, 1325-1333.
- Bartels, A. L., & Leenders, K. L. (2009). Parkinson's disease: The syndrome, the pathogenesis and pathophysiology. *Cortex, 45*, 915-921.
- Blocket, T., Elmar, N., Stemmer, G., Ruzickova, H., & Rusz, J. (2011). Detection of Persons with Parkinson's Disease by Acoustic, Vocal, and Prosodic Analysis. *IEEE*, 5,11
- Braak, Heiko, Del Tredici, Kelly, Rub, Udo, De Vos, Rob A.I., Jansen Steur, Ernst N.H., Braak, Eva. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*, 24, 197-211.
- Braak, Heiko. Ghebremedhin, Estifanos. Rub, Udo. Bratzke, Hansjurgen. Del Tredici, Kelly. (2004). Stages in the development of Parkinson's diseaserelated pathology. *Cell Tissue Research*, 318, 121-134.
- Camacho, A. (2012, July). On the use of auditory models elements to enhance a sawtooth waveform inspired pitch estimator on signals missing low-order harmonics. Paper presented at Proceedings of the 11th International Conference on Information Science, Signal Processing, and their Applications: ISSPA, Montreal, Quebec, Canada (1107-1112).
- Davis, S.B., & Mermelstein, P. (1980). Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28, 357-366.
- Darley, F.L., Aronson, A.E., & Brown, J.R. (1969) *Clusters of Deviant Speech Dimensions in the Dysarthrias.* Rochester, MN: Mayo Clinic and Mayo Foundation.
- Deller, Jr., J.R., Hansen, J.H.L., & Proakis, J.G. (2000). *Discrete-Time Processing of Speech Signals*. New York, NY: IEEE Press.
- Denes, P.B, & Pinson, E.N. (1993) The Speech Chain: The Physics and Biology of Spoken Language Second Edition.. New York, NY: W.H. Freeman and Company

Hawkes, Christopher H., & Deeb, Jacquie. (2006). Predicting Parkinson's disease:

worthwhile but are we there yet? Practical Neurology, 6, 272-277.

- Hillenbrand, J., Cleveland, R.A., & Erickson, R.L. (1994). Acoustic Correlates of Breathy Vocal Quality. *Journal of Speech and Hearing Research*, 37, 769-778.
- Hunter, E.J., & Banks, R. (2014). Tracking Age-Related Speech Characteristics for More Than Forty Years. Poster presented at the 75<sup>th</sup> Annual Michigan Speech Language Hearing Association Conference: MSHA, Kalamazoo, MI.
- Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. Journal of Neurology Neurosurgery and Psychiatry, 79, 368-376.
- Kent, R.D., & Read, C.R. (2002). Acoustic Analysis of Speech: Second Edition. Albany, NY: Thomson Learning Inc.
- Kent, R.D., Weismer, G., Kent, J.F., Vorperian, H.K., & Duffy, J.R. (1999). Acoustic Studies of Dysarthric Speech: Methods, Progress, and Potential. *Journal of Communication Disorders*, 32, 141-186.
- Lansberger, H.A. (1959). Hawthorne Revisited: Management and the Worker, Its Critics, and Developments in Human Relations in Industry. *American* Sociological Review, 24, 277-278.
- Little.,M.A., McSharry, P.E., Hunter, E.J., Spielman, J., & Ramig, L.O. (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *IEEE Transactions for Biomedical Engineering*, 56, 1-9.
- McNeil, Malcolm R., (2009). *Clinical Management of Sensorimotor Speech Disorders*. New York, NY: Thieme Medical Publisher, Inc.
- McReynolds, L.V., & Thompson, C.K. (1986). Flexibility of Single-Subject Experimental Designs. Part I: Review of the Basics of Single-Subject Designs. Journal of Speech and Hearing Disorders, 51, 194-203.
- Parkinson's Disease Foundation (2014). Statistics on Parkinson's Retrieved from http://www.pdf.org/en/parkinson\_statistics on 2/13/2014
- Parkinson's Disease Foundation (2014). Progression Retrieved from http://www.pdf.org/en/progression\_parkinsons on 2/13/2014
- Ramig, Lorraine O., Fox, Cynthia, & Sapir, Shimon. (2008). Speech treatment for Parkinson's disease. *Neurotherapeutics*, 8, 299-311.

- Ramig, Lorraine, A., & Ringel, Robert L., (1983). Effect of Physiological Aging on Selected Acoustic Characteristics of Voice. Journal of Speech, Language, and Hearing Research, 26, 22-30.
- Sadogopan, N., & Smith, A. (2013). Age Differences in Speech Motor Performance on a Novel Speech Task. *Journal of Speech, Language, and Hearing Research*, 56, 1552-1566.
- 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*, 59-88.
- Skodda, S., Gronheit, W., & Schegel, U. (2012). Impairment of Vowel Articulation as a Possible Marker of Disease Progression in Parkinson's Disease. *PLoS ONE* 7 (2): e32132. doi: 10.1371/journal.pone.0032132
- Skowronski, M.D., & Harris, J.G. (2004). Exploiting independent filter bandwidth of human factor cepstral coefficients in automatic speech recognition. *Journal* of the Acoustical Society of America, 116, 3, 1774-1780.
- Skowronski, M.D., Shrivastav, R., Harnsberger, J., Anand, S., and Rosenbek, J., (2012), "Acoustic discrimination of Parkinsonian speech using cepstral measures of articulation," *Journal of the Acoustical. Society of America*, 132, 2089.
- Titze, I.R., & Verdolini Abbott, K. (2012). *Vocology: The Science and Practice of Voice Habilitation.* Salt Lake City, Utah: National Center for Voice and Speech.
- Watts, C.R., & Awan, S.N. (2011). Use of Spectral/Cepstral Analyses for Differentiating Normal from Hypofunctional Voices in Sustained Vowel and Continuous Speech Contexts. *Journal of Speech Language and Hearing Research*, 54, 1525-1537.