An Evaluation of Vocal Analysis Research Relating to Parkinson's Disease

By

Adriana Thompson

_____

A Thesis Submitted to The W.A. Franke Honors College

In Partial Fulfillment of the Bachelors degree
With Honors in

Neuroscience and Cognitive Science

THE UNIVERSITY OF ARIZONA

M A Y 2 0 2 4

Approved by:

Dr. Julie E. Miller
Department of Neuroscience and Speech, Language and Hearing Sciences

Abstract:

Parkinson's Disease (PD) is a progressive neurodegenerative disease affecting almost 10 million individuals globally. It is characterized by tremors, slowness of movement, and vocal dysfunction. Voice dysfunction is one of the earliest indicators of PD and includes breathiness, roughness, reduced loudness, etc. Because most research relates to vocal fold pathology, our understanding of the brain pathology is not well understood prompting insight from animal models such as finch and rodents. These studies have shed light on how the accumulation of Lewy bodies (the aggregated form of the alpha-synuclein protein), and neurites can be key to the progression of symptoms and how it might relate to the vocal measures used in humans. Here, we review current research on the various acoustic measures utilized for PD, evaluate the limitations and strengths of the numerous measures, and provide studies which utilize the same measures. There is a multitude of both speech and voice measures that have been investigated for PD. Various studies demonstrate the potential of these measures in differentiating between those with PD and healthy controls. Yet, there are limitations to current studies including small sample size, methodological inconsistencies, and most importantly, the need for more shared measures. Ultimately, there is a need for further exploration across larger and more diverse subject populations in the future as it will expand our understanding of the disease's pathology, leading to breakthroughs in both PD diagnosis and treatment.

Acknowledgements:

This project was only made possible by the support and generosity of many different people. The mentorship I received on this project from Dr. Julie E. Miller has been invaluable. Thank you for the immense amount of time that was set aside to guide this project. Dr. Robin Samlan, contributed to the editing of both my poster and thesis paper. Thank you for spending valuable time on my project! I could not have completed this without the moral and financial support provided by my parents and would not be graduating this year without them. Finally, thank you to all my family and friends that have supported me through the challenges and celebrations that came with this project.

Chapter 1: Introduction

Background

Parkinson's Disease (PD) is a progressive neurodegenerative disease of the central nervous system. It leads to a wide array of symptoms but predominately affects motor skills and is characterized by muscle tremors, stiffness, slowness of movement, imbalance when walking, etc. About 1.5% of individuals who are 60 or older are diagnosed with PD and on a worldwide level, almost 10 million people suffer from the disease (Sapir, 2014). The progression of the symptoms in those who do suffer from PD is crucial when it comes to both research and diagnosis. Notably, vocal dysfunction is a prevalent concern as between 70 and 90% of individuals with PD develop significant impairments in their voice, and emerging evidence has suggested that these voice abnormalities are one of the earliest indicators of the disease as vocalization is very complex and involves very fine motor control (Ma, 2019). These vocal impairments are characterized by breathiness, roughness, and reduced loudness. Most people with these dysfunctions have a disorder known as hypokinetic dysarthria (HKD). Those who suffer from HKD tend to have monopitch, monoloudness, imprecise consonants, variable rate, and breathy vocal quality with the most distinctive symptoms being reduced loudness, known as hypophonia, and a monotone speech pattern (Sapir, 2014). These extensive speech abnormalities are accompanied by other symptoms such as swallowing issues along with quality of life issues. We do not know how the PD pathology affects vocal changes.

Brain Pathology:

While it is true that the neuropathological mechanisms underlying voice abnormalities of PD are not well known, new studies have begun to shed light on how the accumulation of Lewy bodies, which consist largely of aggregated forms of the alpha-synuclein protein, and neurites, referred to as Lewy pathology, is key to the development of PD symptoms. The pathology first starts in the anterior olfactory nucleus and the lower part of the brain, known as the brainstem (Ma et al., 2019). The pathology then spreads progressively in multiple stages and from this progression, we can infer possible origins of voice dysfunction in PD.

In the early stages, the dorsal motor nucleus of the vagus and glossopharyngeal nerves in the medulla are affected. The vagus nerve provides input to the laryngeal muscles and is necessary for both swallowing and speech while the glossopharyngeal nerve innervates the stylopharyngeus muscle and is responsible for elevating the larynx and pharynx (Thomas, 2022). The spread to this area is a precursor to the spread of the pathology into more basal ganglia and cortical areas that impact speech production (Ma et al., 2019). Lewy pathology is eventually found in numerous areas of the pons within the brainstem, which are critical when it comes to basic emotional processing (as nuanced control of emotions is primarily governed by higher cortical regions) and expressing those emotions through speech and voice. Dysfunction in these areas can lead to dysprosody (speech that has an absent or atypical rhythm, intonation, or melody). Researchers have found that speech impairments in PD is greater in emotional situations than in neutral ones. Researchers have also found specific brainstem areas that regulate our emotions through using various imaging and pathological studies. These areas include the raphe nuclei, associated with serotonin;

the locus coeruleus, which uses noradrenaline; the ventral tegmental area, linked with dopamine; and the pedunculopontine and laterodorsal tegmental nuclei, which are both influenced by acetylcholine. It has been found that Lewy bodies extensively infiltrate nearly all these areas in PD. The locus coeruleus is particularly notable because it has significantly more neuronal loss than even in the substantia nigra, which is commonly associated with loss of dopamine in PD (Ma et al., 2019). This pronounced degradation in the locus coeruleus may be the reason for why speech rhythm and tone issues are more apparent in PD's early phases than problems with voice strength or clear articulation.

This is consistent with studies showing that 60.87% of individuals in the early stages of PD experienced speech rhythm and tone difficulties, while issues with articulation and voice strength were less frequent (Ma et al., 2019). Articulation refers to how clearly someone forms words and involves the tongue, teeth, jaw, and lips. Alternatively, research into another condition, REM sleep behavior disorder, which potentially leads to PD, primarily identified articulation and voice strength problems, indicating that different patterns of brain damage may distinctly influence voice and speech (Ma et al., 2019).

Eventually the pathology spreads to the temporal mesocortex and to the adjacent higher-order sensory association areas within the mature neocortex (Braak, 2003). By the next stage, the spread of the cortical pathology becomes even more extensive, affecting first-order sensory association areas, premotor areas, and sometimes even primary sensory and motor areas. Within the temporal lobe, specifically, the pathology extends into the cortex of the first temporal convolution and often to the primary auditory area. The spread and significant damage to areas such as the amygdala, hippocampal

formation, and anteromedial temporal mesocortex alongside the damage to the neocortical regions is thought to be a major contributor to the significant cognitive decline in PD (Braak, 2003).

By the later stages of PD, Lewy bodies cause harm to the dopamine-producing neurons in the substantia nigra resulting in a loss of dopamine input to the basal ganglia. This loss of cells results in dopamine deficiency, causing stiffness and diminished movements in the muscles that are responsible for speech and breathing, contributing to the various speech issues in PD. Ultimately, dopamine cell death occurs relatively mid to late stages in the course of the disease, while brainstem motor systems are likely to be affected at the earlier stages of PD and may involve non-dopaminergic mechanisms (Sapir, 2014).

We can infer how various brain changes affect and map onto vocal impairments in PD but most of our understanding relating to brain pathology comes from animal models. In order to understand the progression and underlying causes of communication challenges in PD, a variety of rat and mouse models have been found to be profoundly useful (Krasko, 2021). Some examples include neurotoxin approaches using a dopamine-depleting drug, or incorporating genetic mutations in known human PD-associated genes (*SNCA*-alpha-synuclein, *PINK-1*). Each PD model offers its own unique insights and also has its drawbacks. Neurotoxin models, for instance, are useful for studying the later stages of PD movement problems associated with the dopamine loss, but they lack the earlier events in the disease including the buildup of toxic forms of alpha-synuclein protein (Krasko, 2021). Songbirds offer a unique model for understanding both the contributions of dopamine and alpha-synuclein protein to the

vocal output (birdsong) given their song-dedicated brain nuclei (Miller et al., 2015; Medina et al. 2022). Ultimately, while we do not know specifically how brain changes map onto voice and speech symptoms, there is extensive knowledge about changes in vocal fold physiology for individuals that do suffer from vocal impairments in PD.

Vocal Physiology:

Human vocalizations are generated through the integration of multiple processes with a large role played by the respiratory system. It begins with air moving out of the lungs (Ma et al., 2019). Phonation, or the production of speech sounds, results from vibrations at the vocal folds as a result of this air flow. Subsequently, this sound is shaped by structures including the lips, teeth, and tongue to produce speech. The vocal folds form a valve that functions for breathing, swallowing, and speaking. When the vocal folds adduct, they both protect the lower airway (trachea and lungs) and place the vocal folds in position to vibrate for talking. There are three main cartilages important to vocal fold movement; the singular thyroid, cricoid, and a pair of arytenoid cartilages. The thyroid cartilage is shield-like and rests on top of the ring-shaped cricoid cartilage. The arytenoid cartilage rests on the cricoid and has pyramid-like shape to it. The vocal folds attach anteriorly to the thyroid cartilage and posteriorly to the two arytenoid cartilages. The cartilages are crucial for sound production because they provide attachment points for the muscles that alter the position and tension of the vocal folds. The intrinsic laryngeal muscles, specifically, abduct and adduct the vocal folds in addition to influencing their length and tension. The posterior cricoarytenoid muscle is responsible for abduction and opens the airway by pivoting the arytenoid cartilages outward to separate them. There are multiple muscles involved in adduction including the

thyroarytenoid, lateral cricoarytenoid, and interarytenoid muscles. There is a section in the thyroarytenoid muscle that connects to the vocal process known as the vocalis muscle. This muscle works to shorten and thicken the vocal folds, reducing the pitch. Alternatively, the cricothyroid muscle pulls the anterior thyroid cartilage towards the anterior cricoid cartilage, stretching and tightening the vocal fold, and leads to an increase in pitch.

These muscles are innervated by branches from the vagus nerve (Ma et al., 2019). All of the intrinsic laryngeal muscles are innervated by the recurrent laryngeal nerve, with the exception of the cricothyroid muscle as it is supplied by the external branch of the superior laryngeal nerve. The vagus nerve stems out of several nuclei of the medulla, with the neurons of the intrinsic laryngeal muscles specifically coming from the nucleus ambiguous. At its core, voice is produced by the motor and sensory nuclei in the lower brainstem and spinal cord, which are connected by the lateral reticular formation nuclei (Ma et al., 2019). Together, these nuclei control laryngeal and articulatory functions. There is much less understanding of higher order neural control of vocalization. However, recent models propose two different functional pathways with one network thought to control voluntary vocalization and a second network to control innate emotional and non-verbal vocalizations such as laughing and crying.

Parkinson's Impact on the Vocal Folds:

There have been numerous techniques utilized to identify the effects of PD on vocalization (Ma et al., 2019). Laryngoscopy, for example, gives direct visualization of the vocal folds and shows bowed vocal fold edges and incomplete glottal closure in

individuals with PD because of weakened movement. The incomplete closure can cause air leakage during speech, diminishing voice quality and resulting in breathiness, weakness, and reduced voice loudness. Research has also indicated that individuals with advanced PD have asymmetry in their vocal fold closure with more bowing on the side of the body most affected by the disease. Additionally, electromyography has revealed that patients with PD have increased muscle activity at rest which can lead to severe speech impairments due to a loss in normal muscle coordination. Respiratory assessment has also shown that the generation of airflow necessary for vocalization can be compromised in individuals with PD. Aerodynamic studies have shown reductions in subglottal pressure, laryngeal resistance, and peak airflow due to restrictive respiratory dysfunction and reductions in maximal inspiratory and expiratory pressures (Ma et al., 2019).

Ultimately, Vocal dysfunction is a significant and early indicator of PD and is known to lead to breathiness, roughness, and reduced loudness. The underlying brain pathology of these symptoms involves the accumulation of Lewy bodies and neurites in the brainstem and cortical regions, affecting vocal function. This can potentially lead to vocal (voice + speech) issues early on in the disease. Our understanding of this PD pathology is continuously evolving and insights from animal models such as songbirds and rodents have been immensely valuable. Additionally, understanding human vocalization is necessary to further PD research and to assess the contributions of dopamine loss and toxic forms of the alpha-synuclein protein.

In Chapter 2, I will be presenting the current voice and speech measures researchers are utilizing in human subject and animal models of PD.

References

Braak, H., Del Tredici, K., 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.

Krasko M.D., Hoffmeister J.D., Schaen-Heacock N.E., Welsch J.M., Kelm-Nelson C.A., Ciucci M.R. Rat Models Of Vocal Deficits in Parkinson's Disease. Brain Sciences. 2021; 11(7):925.

Ma, A., Lau, K. K., & Thyagarajan, D. (2020). Voice Changes in Parkinson's Disease: What are They Telling Us?. Journal of clinical neuroscience: official journal of the Neurosurgical Society of Australasia, 72, 1–7.

Medina, C. A., Vargas, E., Munger, S. J., & Miller, J. E. (2022). Vocal Changes in a Zebra Finch Model of Parkinson's Disease Characterized by Alpha-synuclein Overexpression in the Song-dedicated Anterior Forebrain Pathway. PloS one, 17(5).

Miller, J. E., Hafzalla, G. W., Burkett, Z. D., Fox, C. M., & White, S. A. (2015). Reduced Vocal Variability in a Zebra Finch Model of Dopamine Depletion: Implications for Parkinson Disease. Physiological reports, 3(11).

Sapir S. (2014). Multiple factors are Involved in the Dysarthria Associated with Parkinson's Disease: a Review with Implications for Clinical Practice and Research. Journal of speech, language, and hearing research: JSLHR, 57(4), 1330–1343.

Chapter 2: Current Approaches to Measuring Acoustic Features of Vocalizations Across Humans and Zebra Finches

There is a multitude of acoustic measures that researchers utilize to further the study of PD. This chapter will dive into the variety of measures that have recently been explored and compare between the studies in their ability to move the field forward.

In my review of the primary literature, I observed that there was not a lot of overlap in measurements used; therefore, I first discuss each paper individually before consideration of common themes. An overview of the various acoustic measures is provided in this table below.

Voice or Speech Measure	Perceived As	Units	Description	Human or Finch Measure
Intensity/Amplitude[1,2,3]	loudness	dB SPL	amount or intensity of sound; volume	Both
Fundamental Frequency[1,2,3]	pitch	Hertz	how high or low a sound is	Both
Cepstral Peak Prominence (CPPS)[2,3] and Standard Deviation (SD)[4]	voice quality (in humans)	dB	describes the prominence of periodic (i.e., harmonic) energy in the acoustic waveform standard deviation of CPP measures across sample	Both

Cepstral/Spectral Index of Dysphonia (CSID)[4]	voice quality (dysphonia)	dB	computation involving CPP, CPPSD, and L/H SR with different weightings for speech material and sex	Human
Mel Frequency Cepstral Coefficient (MFCC)[1,5]		unit of time	1.quantifies vocal fold dynamics depending on properties of the articulators 2. used for catching the PD effects in vocal tract separately from the vocal folds	Human
Wiener Entropy (WE)[2,3]	quality (in finch)	measured on a logarithmic scale from zero to minus infinity.	a measure of the periodic versus aperiodic energy in a birdsong syllable	Finch
Jitter[1,6]	voice quality	percentage (%)	the parameter to measure frequency changes from cycle to cycle	Human
Shimmer[1,6]	roughness	dB, %	measuring cycle to cycle changes in the amplitude of the sound wave	Human

GNE (Glottis to Noise Excitation)[1,5]	extent of turbulent noise or breathiness		degree of signal strength vs noise	Human
HNR (Harmonic to Noise ratio)[3,5,7]	voice quality	dB	ratio of harmonic (i.e., periodic) energy to inharmonic (i.e., noise) energy	Both
L/H Spectral Ratio (L/H SR)[4]	voice quality	dB	mean ratio of energy below 4 kHz to energy above 4 kHz.	Human
Consonant Cluster Production[8]	intelligibility	spectrogram ratings on six-point scale	utilized six-point rating scale, 0= atypical, to 5=typical	Human
Alternating Motion Rate (AMR)[8]		syllables/sec	repeating the same syllable as fast as possible	Human
Sequential Motion Rate (SMR)[8]		syllables/sec	repeating a sequence of different syllables	Human
Voice Onset Time (VOT)[8]	whether stop consonants are perceived as voiced or voiceless	time (ms)	measures the time between the release of the oral constriction for plosive (stop) production and the onset of periodic vocal fold vibration	Human

As shown in the table above, both speech and vocal measures have been investigated to research PD. Additionally, some measures (such as WE) have only been applied to finch while others (Jitter) have only been applied to humans. It is important to note, however, that measures such as fundamental frequency and intensity have been used across both birdsong and human voice. [1]Altay, Medical Hypotheses, 2020. [2]Badwal et al., Journal of speech, language, and hearing research, 2019. [3]Badwal et al. 2020. [4]Alharbi et al.,Folia phoniatrica et logopaedica (IALP), 2019. [5]Arora et al., Journal of Parkinson's Disease, 2018. [6]Azadi et al., Adv Biomed Res, 2021. [7]Ma et al., Journal of Clinical Neuroscience, 2020. [8]Johansson et al., Clinical Linguistics and Phonetics, 2023.

Utilizing Voice to Discriminate Participants with *LRRK2*-associated PD from Idiopathic PD (iPD):

Past work has demonstrated that objective measurement of vocal impairments can be utilized to distinguish iPD participants from healthy controls with high accuracy. The term Idiopathic in iPD indicates that the cause is unknown. Additionally, "voice-based measures have been used to accurately replicate both the motor and total Unified Parkinson's Disease Rating Scale assessment" (Arora et al., 2018). The Unified Parkinson's Disease Rating Scale assessment is the most widely used PD rating scale and assesses different aspects of PD (mentation, behavior and mood, activities of daily living, and motor examination). Ultimately, it evaluates the prevalence of different characteristics of PD including intellectual impairment, depression, speech, swallowing, handwriting, etc. These findings encouraged Arora et al to further the investigation of voice analysis as a reliable non-invasive tool for identifying PD. Since *LRRK2* mutations are one of the most common causes of genetically determined PD (along with

SNCA, *PINK1*, *PARK7*, etc), identifying the prodromal state for early intervention is crucial. In their study, they analyzed voice-based measures in families carrying a mutation in the gene for *LRRK2*. They had two different goals from this study. They aimed to determine if voice could be utilized to discriminate participants with *LRRK2*-associated PD from idiopathic PD. They also wanted to examine if there are any differences in voice between non-manifesting carriers (NMC) of a *LRRK2* mutation, non-manifesting non-mutation carriers, and unrelated healthy controls.

They evaluated the presence or absence of a *LRRK2* mutation in all of the participants and obtained two audio recordings of sustained vocal phonation from all the participants to differentiate the groups by analyzing salient voice features (seven participants with *LRRK2*-PD, 17 participants with iPD, 20 non-manifesting carriers (NMC) of *LRRK2* Mutations, 25 related non-manifesting non-carriers, and 26 healthy controls were recruited). The study conducted paired comparisons across three groups: *LRRK2*-PD versus iPD (idiopathic PD), NMC versus related non-carriers (RNC), and NMC versus healthy controls. In the *LRRK2*-PD versus iPD group, Glottis to Noise Excitation (GNE) was analyzed and revealed significant differences which indicated that it is highly discriminatory between *LRRK2*-PD and iPD. The study also examined Harmonic to Noise Ratio (HNR) and shimmer in the NMC versus RNC group. HNR revealed a significant difference, suggesting a discriminative ability between the two while shimmer did not. Lastly, Mel Frequency Cepstral Coefficient (MFCC) and Shimmer were analyzed in the NMC versus healthy controls group. Both MFCC and shimmer were significantly different between NMC and healthy controls. They also

compared UPDRSIII among the carriers and controls and found that NMC had a higher UPDRSIII score compared to healthy controls and related non carriers.

Ultimately, the researchers found significant differences between *LRRK2*-PD and iPD in the sustained phonations features. These differences in features were less pronounced when comparing the NMC, RMC, and healthy control groups. Thus, voice could potentially become a non-invasive biomarker for identifying a *LRRK2* mutation in PD participants. The study does warrant further investigation into utilizing voice as an early biomarker but did have some limitations. The researchers had a relatively small sample size, especially those with *LRRK2*-PD and there were also some inconsistencies in the study as the patients with iPD were older than those with *LRRK2*-PD.

The Effect of PD on Additional Voice/Speech Features

Similar to Arora et al, Azadi et al (2021) also utilize shimmer with the goal of making voice analysis a method for initial diagnosis of PD. There have been many methods for analyzing the change in sound. However, many different factors such as the environment or the quality of the microphones have become an issue when analyzing the data. It is also difficult to evaluate large numbers of acoustic parameters. Thus, it is critical to find the minimal set of features that are optimal in providing the most reliable diagnosis. The researchers in this study utilized the measurements jitter and shimmer. The study includes four main stages which are “data acquisition, feature extraction, feature selection, and evaluation of the classifier performance in noise-free and noisy conditions.” Their study compares healthy individuals with people who have PD. The samples were taken from an elderly care center and a clinical office of neurologists.

Those with PD were also divided into mild, moderate, and severe. The vowel /a/ was then collected from each participant (several voice samples were collected). They then added disturbances to the signal to simulate noise on telephone lines. This way, an initial assessment can be made. They extracted features from the noise-free and noisy signals separately to find which features perform better.

They then evaluated the signals with the jitter and shimmer measurements. As mentioned in the table above, jitter is the parameter to measure frequency changes from cycle to cycle and shimmer is related to measuring changes in the amplitude speech wave. They found that when comparing healthy men with men who have PD, the values of both jitter and shimmer increased. However, the values did the opposite for women as the values of the extracted features decreased compared to the healthy group. The researchers also found, according to an LS-SVM classifier (Least squares SVM classifiers are “a class of kernel-based learning methods to solve both classification and regression problems”) that the features had about a 70% accuracy in separating people with PD from healthy individuals in noiseless and noisy conditions for each gender. When both features were used together, the accuracy increased to about 80%.

Their findings potentially move the field forward as they provide a method that could distinguish patients with PD from healthy individuals with a minimal set of voice features. Specialists would be able to screen PWP remotely and people could be diagnosed early on. However, it would’ve been beneficial if the researchers touched more on the different results between the participants that had mild PD versus those who had severe PD and whether the accuracy of differentiating those with PD from

healthy individuals changed based on the individual's stage of PD as individuals with mild PD would be the crucial ones for early diagnosis and treatment intervention).

Evaluation of Utilizing Shimmer to Differentiate PD from Healthy Individuals:

Both of the studies above focused on the potential of voice analysis for distinguishing between individuals with PD and healthy individuals. However, they did so by utilizing different approaches and measures. Arora et al.'s study focused on the genetic aspects of PD, specifically pertaining to the *LRRK-2* mutation, suggesting voice analysis could serve as a biomarker for genetically determined PD. Alternatively, Azadi et al. offered a broader approach, with the genetic status of the individual with PD unknown, that could potentially lead to early diagnosis. Both studies, however, did utilize the measure shimmer and found it to be a promising measure as an early indicator of PD. Both also highlighted the need for larger samples and more consistent data in the participants with PD. Arora et al.'s study, specifically, had older participants with iPD than those with *LRRK2*-PD.

The Utilization of Speech Measures and Their Relation to Intelligibility

It is apparent from the studies above that there are bodies of research on acoustic aspects of PD. However, research relating acoustic measures to intelligibility is still very limited, and there is the need for more studies on dysarthria in individuals who are not English speakers in order to expand the cross-linguistic knowledge base. Thus, Johansson et al. (2023) conducted a study aiming to explore acoustic measures of speech production in Swedish speakers with PD. They specifically focus on consonant cluster production, articulatory motion rate and variation (AMR), and voice onset time

(VOT), all of which are described in the table above. The researchers ultimately wanted to investigate how these specific acoustic measures relate to speech intelligibility in Swedish speakers with PD.

The researchers conducted a study with eleven speakers experiencing speech problems due to PD and six neurologically healthy individuals. The distribution for age, gender, and regional variants of Swedish were similar for the two groups. Dysarthria was assessed through the Swedish Dysarthria Assessment and was reported to be mild, moderate, or severe with four, five, and two in each category, respectively. All of the participants were recorded in quiet environments and repeated alternate and sequential movements where they repeated the “syllables /pa/, /ta/, /ka/, and the syllable sequence /pataka/ at maximum rate” for five seconds. They also participated in oral reading of “phonetically balanced sentences from the Swedish dysarthria assessment.” Their intelligibility was then measured. For statistical analysis, the study combined ratings from evaluators to derive average values and used non-parametric statistical methods to compare the acoustic measurements between people with PD and people without. The study revealed that at a group level, ratings of consonant clusters were lower for people with PD than for those without (healthy speakers). The difference was statistically significant for the clusters. For individual speech sounds in clusters, the ratings at group level were much lower for speakers with PD but there was considerable variation of individual speech sounds in clusters when comparing speakers with PD to healthy speakers. Additionally, at a group level, AMR (alternating motion rate) was much slower for repeated lingual movements for people with PD compared to healthy speakers. AMR variability was also significantly greater for PD

speakers in all three articulatory positions. Individual results, however, showed no clear relationship with the speakers' levels of intelligibility. VOT (voice onset time) was also significantly longer for PD speakers than for healthy speakers in all articulatory positions. The strongest correlation between intelligibility and acoustic measures was that speakers that had lower intelligibility tended to have lower consonant cluster ratings than those with higher intelligibility.

This study gives insight on the possible correlations between acoustic measures and intelligibility. This is crucial knowledge as it can lead to the idea that acoustic measures can be utilized for early diagnosis of PD. However, this study did have a small sample size and the results were difficult to interpret as there may be at least two different ways to interpret the relationship between acoustic measures and intelligibility. One is that a specific acoustic variable reflects the overall level of impairment, and is just one affected part among others, which together affect intelligibility. Another interpretation suggested by Weismer et al. (2001) is that a specific acoustic variable directly influences intelligibility. "The strong correlations between the consonant cluster ratings and intelligibility stand out and raise the question of whether consonant cluster articulation has a particularly strong influence on intelligibility."

The Use of CPP in PD Diagnosis and Treatment

Alharbi et al. (2019) also utilized CPP to investigate how voice treatment can alter the results of voice measures. Intensive behavioral voice treatment such as the Lee Silverman Voice Treatment (LSVT) has been proven to be very effective when compared to pharmacologic or surgical alternatives (Alharbi et al, 2019). The treatment

has been found to improve speech in numerous ways such as increasing fundamental frequency range, fundamental frequency variability in speech, and increasing vocal intensity. It can also result in changes in physiology as vocal fold closure has been shown to improve. However, there have been no studies that examine spectral/cepstral measure of dysphonia before and after the treatment. As described in the table, CPPS essentially assesses the degree to which harmonics dominate the spectrum of the voice signal. The low/high spectral ratio has also shown to be effective in assessing different types of dysphonic voices in sustained vowels and continuous speech. The Cepstral Spectral Index of Dysphonia (CSID) incorporates these two measures above in specific combinations that are "differentially suited for sustained vowels or connected speech sampling." The study in this paper expands upon previous work which found there to be physiological changes. Thus, Alharbi et al. predicts that there should also be increased loudness following LSVT along with changes in the spectrum, specifically in harmonic structure. Evidence from previous studies suggest that as voice intensity increases following LSVT, there should also be a systemic increase in the dominance of harmonics in people with PD (Alharbi et al., 2019).

Alharbi et al.'s goals in this study were to measure the changes in CPP, CPP standard deviation, L/H SR, and CSID before and after LSVT in order to enhance our understanding of how treatment can alter the results of these measurements. The researchers emphasized that the purpose of their study was to utilize a naturalistic method to cause cepstral/spectral changes correlated with increases in voice intensity and vocal fold closure rather than to evaluate LSVT. The researchers hypothesized that the measurements would correlate to stronger harmonic structure after treatment with a

lower CPP SD and higher CPP, leading overall to decreased dysphonia, measured by cepstral/spectral index of dysphonia (CSID). To test this hypothesis, nine adults were included in the study with six males and three females ranging between 52 and 81 years. All of the participants were evaluated to verify the presence of hypokinetic dysarthria with hypophonia; and each participant underwent standard LSVT administered by an ASHA-certified speech-language pathologist. All speakers enrolled in LSVT participated in a total of 16 sessions for 4 days a week for 4 weeks. Voice recordings were obtained on 3 different days within 1 week before and on 3 different days within 1 week after LSVT. Each speaker was instructed to produce 3 repetitions of a sustained vowel /a/ holding out the vowel as steady and as long as she or he could. For each vowel recording, CPP, CPP SD, CSID were observed through a ADSV program. L/H SR was analyzed using two different frequency cutoffs.

The study demonstrated that the participants' sustained vowels increased intensity from pre-treatment to post-treatment. ANOVA results also showed that four acoustic variables demonstrated statistically significant differences from pre- to post-LSVT including CPP, CPP SD, L/H SR cutoff, and CSID. Specifically, CPP and CPP SD increased significantly while CSID decreased significantly. The findings of this study reveal that the measures which were utilized by Alharbi et al. are associated with increased vocal intensity. Speakers with PD exhibited both improved harmonic structure and voice quality which was reflected by cepstral/spectral analysis. This will help move the field forward as it is important to find which measures can accurately detect improvements in speech with treatment and how analyzing harmonic structure can be utilized in further research relating to the disease. However, this study was utilized with

participants that demonstrated rather severe hypokinetic dysarthria so it would be beneficial to see the results for those who are in the beginning stages of the disease.

Comparative Analysis Between Voice Measures in Human Populations with Birdsong:

As mentioned in the introduction, much of our understanding of the brain pathology of PD originates from animal models. The data obtained from animal models of PD can be applied to humans particularly the adult male zebra finch songbird (Sakata et al., 2020). There are anatomical and genetic similarities between finch brain nuclei regions dedicated to song learning and production with human brain regions for vocal learning and production including a cortico-basal ganglia-thalamo-cortical loop that is under neuromodulatory control by dopamine (Pfenning et al, 2014; Simonyan et al., 2012).

Research studies that have been conducted using songbirds utilize and rely on a standard set of acoustic measures using their specific software programs to describe the birdsong (Badwal et al., 2019). Many of these measurements are unfamiliar to the researchers who are working with human voice, and vice versa, hindering their ability to exchange information. Thus, a “common language” between birdsong and human voice needs to be introduced. The Badwal et al. study (2019) had several goals. The first was to create a “translational dictionary" that could connect the measurements of both birdsong and humans to allow collaborative research. The second was to see if Wiener Entropy (WE) used in birdsong and Cepstral Peak Prominence (CPPS) could provide similar information when applied to human voice data. The basic unit of birdsong is a sequence of repeated syllables known as a motif. A motif is encoded by specific

neuronal firing patterns in song-dedicated brain nuclei. The syllables are defined by their spectral profile and are labeled as either “noisy,” “harmonic,” or “mixed.” Mixed syllables are composed of both harmonic and a noisy note. For the song analyses, three adult male zebra finch that were at the midpoint of their lifespan were used (Badwal et al., 2019). The researchers hypothesized that the human voice measure CPPS and the birdsong measure WE would provide the same information about the harmonic and noise components of the birdsong syllable.

However, they found that the relationship between CPPS and WE is complex (Badwal et al., 2019). For the syllables that were clearly noisy or harmonic, their CPPS and WE scores showed a consistently inverse relationship. However, for the mixed bird syllables, this relationship was not preserved because the WE scores were too variable. Thus, for mixed syllables, CPPS scores appeared to be more reliable. However, there were also instances when CPPS was similar for both noisy and harmonic syllables while WE score was able to differentiate the syllables. A study from the same group showed that age-related changes in birdsong syllables could be detected using CPP and WE (Badwal et al. 2020). Ultimately, these findings allow for further development of comparative analyses between voice measures in human populations with birdsong. This is crucial for being able to compare the information obtained from laboratory animal models with diagnosis and treatment of human neurodegenerative diseases.

Use of CPP in Human Voice Studies vs. Birdsong:

The two studies discussed above both utilize CPP as a measure in their analyses. However, the contexts and the objectives of their studies differed significantly. Badwal

et al. aimed to find a common ground between human and animal measures to advance PD research, but examined normal aging finches. Alharbi et al. on the other hand, focused on how measures are affected by improved speech functions in PD patients due to a targeted behavioral treatment. Despite their distinct goals, both studies emphasize the significance of voice analysis in advancing our understanding of PD.

In conclusion, a range of studies have investigated the use of acoustic measures to diagnosis and understand PD. The results from these studies have demonstrated the potential of numerous kinds of vocal and speech measures such as jitter, shimmer, CPP, and MFCC in differentiating between those with PD and healthy controls. They have also illustrated the potential in differentiating different subgroups of PD patients, such as those with genetic mutations like *LRRK2.* Most of these studies also revealed the promising application of voice analysis as a non-invasive diagnostic tool for early detection and intervention of PD. However, these studies did include many limitations such as small sample sizes, methodological inconsistencies, and overall, the need for more research pertaining to the same kinds of acoustic measures.

Future Directions in PD Vocal Research:

Ultimately, further exploration across larger and more diverse populations is necessary in future studies as this will both solidify the role of acoustic measures in PD research and expand our understanding of the disease's pathology, leading to breakthroughs in both PD diagnosis and treatment. While the spread of Lewy pathology is understood to occur in multiple stages, its early and specific impact on vocalization compared to other motor functions is not fully understood. Thus, studying the accumulation of aggregated

forms of the alpha-synuclein protein can provide promising discoveries relating to PD vocal dysfunction. Animal models have continued to deliver immensely valuable findings for the field. Therefore, future use of these models have the potential to help gain more of an understanding of vocal problems in PD as the field evolves, to identify promising new avenues related to non-invasive PD biomarkers.

References

Alharbi, G. G., Cannito, M. P., Buder, E. H., & Awan, S. N. (2019). Spectral/Cepstral Analyses of Phonation in Parkinson's Disease before and after Voice Treatment: A Preliminary Study. Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics (IALP), 71(5-6), 275–285.

Altay, E. V., & Alatas, B. (2020). Association Analysis of Parkinson Disease with Vocal Change Characteristics Using Multi-Objective Metaheuristic Optimization. Medical Hypotheses, 141, 109722.

Arora, S., Visanji, N. P., Mestre, T. A., Tsanas, A., AlDakheel, A., Connolly, B. S., Gasca-Salas, C., Kern, D. S., Jain, J., Slow, E. J., Faust-Socher, A., Lang, A. E., Little, M. A., & Marras, C. (2018). Investigating Voice as a Biomarker for Leucine-Rich Repeat Kinase 2-Associated Parkinson's Disease. Journal of Parkinson's disease, 8(4), 503–510.

Azadi, H., Akbarzadeh-T, M. R., Shoeibi, A., & Kobravi, H. R. (2021). Evaluating the Effect of Parkinson's Disease on Jitter and Shimmer Speech Features. Advanced biomedical research, 10, 54.

Badwal, A., Poertner, J., Samlan, R. A., & Miller, J. E. (2019). Common Terminology and Acoustic Measures for Human Voice and Birdsong. Journal of speech, language, and hearing research: JSLHR, 62(1), 60–69.

Badwal, A., Borgstrom, M., Samlan, R.A., & Miller, J.E. (2020). Middle Age, a Key Timepoint for Changes in Birdsong and Human Voice. Behav Neurosci, 134(3): 208-221.

Johansson, I. L., Samuelsson, C., & Müller, N. (2023). Consonant Articulation Acoustics and Intelligibility in Swedish Speakers with Parkinson's Disease: a pilot study. Clinical linguistics & phonetics, 37(9), 845–865.

Ma, A., Lau, K. K., & Thyagarajan, D. (2020). Voice Changes in Parkinson's Disease: What are They Telling Us?. Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia, 72, 1–7.

Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., Howard, J. T., Wirthlin, M., Lovell, P. V., Ganapathy, G., Mouncastle, J., Moseley, M. A., Thompson, J. W., Soderblom, E. J., Iriki, A., Kato, M., Gilbert, M. T., Zhang, G., Bakken, T., Bongaarts, A., … Jarvis, E. D. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. Science (New York, N.Y.), 346(6215), 1256846.

Sakata, J. T., Woolley, S. C., Fay, R. R. & Popper, A. N. The Neuroethology of Birdsong. Vol. 71 (Springer, 2020).

Simonyan, K., Horwitz, B., & Jarvis, E. D. (2012). Dopamine Regulation of Human Speech and Bird Song: a critical review. Brain and language, 122(3), 142–150.