Louisiana State University
LSU Digital Commons

LSU Master's Theses | Graduate School

2016

Reliability of Auditory-Perceptual Ratings of Dysarthric Speech: Hypokinetic Dysarthria Secondary to Parkinson's Disease

Jessica Lynn Miller
Louisiana State University and Agricultural and Mechanical College, jmil143@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

 Part of the Communication Sciences and Disorders Commons

Recommended Citation
Miller, Jessica Lynn, "Reliability of Auditory-Perceptual Ratings of Dysarthric Speech: Hypokinetic Dysarthria Secondary to Parkinson's Disease" (2016). *LSU Master's Theses*. 2037.
https://digitalcommons.lsu.edu/gradschool_theses/2037

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

RELIABILITY OF AUDITORY-PERCEPTUAL RATINGS OF DYSARTHRIC SPEECH: HYPOKINETIC DYSARTHRIA SECONDARY TO PARKINSON'S DISEASE

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Communication Sciences and Disorders

by
Jessica Lynn Miller
B.A., Louisiana State University, 2013
May 2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Perceptual judgment has been the gold standard in clinical practice, especially regarding differential diagnosis and treatment of dysarthria. Thus, it is critical to establish the reliability of perceptual ratings of the speech characteristics associated with different types of dysarthria. Despite its importance, the reliability and sensitivity of perceptual ratings of speech disturbance have been somewhat questioned. The purpose of this study was to examine the interrater reliability of ratings of perceptual characteristics and the saliency of these characteristics as related to hypokinetic dysarthria. Due to the feasibility issue, the scope of the study was limited to hypokinetic dysarthria associated with Parkinson's disease.

Eight subjects with hypokinetic dysarthia and three healthy controls were selected from a large dysarthria speech database for the study. All dysarthia subjects were diagnosed with Parkinson's disease and dysarthria. Recordings of a standard reading passage were played to 148 raters who were inexperienced undergraduate students majoring in communication sciences and disorders at Louisiana State University. Raters' results were statistically analyzed to determine interrater reliability across 37 dimensions. The mean score value of each dimension was also calculated for the control and Parkinson group and statistically compared to determine the most salient dimensions.

Results demonstrated relatively low intra-rater reliability with 52 listeners of 148 demonstrating reliability above 0.60. Interrater reliability for dimensions ranged from 0.007 to 0.730. Twenty-four dimensions were identified as salient for hypokinetic dysarthria, all of which were statistically different between the control speaker group and the PD speaker group.

INTRODUCTION

Importance of perceptual ratings of dysarthria

Until 1969, the term dysarthria was largely defined as "imperfect articulation in speech" (Dorland, 1965). In recognition of its inappropriateness given that dysarthria exhibits complex speech production abnormalities beyond articulation, Darly, Aronson, and Brown (1969a) redefined dysarthria with a more comprehensive and specific definition as follows: "Dysarthria is a collective name for a group of speech disorders resulting from disturbances in muscular control over the speech mechanism due to damage of the central or peripheral nervous system" (p. 246). This classic study of dysarthria suggested the following six types: flaccid, spastic, ataxic, hyperkinetic (chorea or dysarthria), hypokinetic, and mixed (flaccid-spastic). Each of these classifications is often related to an underlying neurologic condition, as well as the presence of deviant speech dimensions identified by Darley, Aronson, and Brown (1969a; 1969b) (Fawcett, 2010).

In their effort to differentiate among the different types of dysarthrias, Darly et al. (1969a) collected speech samples from patients representing seven neurologic disorders: pseudobulbar palsy, bulbar palsy, amyotrophic lateral sclerosis, cerebellar lesions, parkinsonism, dystonia, and choreoathetosis. The participants represented a wide range of severity of speech involvement. Despite the heterogeneity of the participants in terms of neuropathologies and speech severity, the authors identified 38 perceptual dimensions that deviated in the speech of the dysarthrias, which have provided the basis of dysarthria classification. A description of each dimension is presented in the Appendix A.

Deviant speech dimensions by DAB

The 38 dimensions are grouped into seven categories. Four dimensions pertain to *pitch*: pitch level, pitch breaks, monopitch, and voice tremor. Five dimensions pertain to *loudness*: Monoloudness, excess loudness variation, loudness decay, alternating loudness, and loudness level overall. Nine dimensions pertain to *vocal quality*: harsh voice, hoarse (wet) voice, breathy voice (continuous), breathy voice (transient), strained-strangled voice, voice stoppages, hypernasality, hyponasality, and nasal emission. Three dimensions pertain to *respiration*: forced inspiration-expiration, audible inspiration, and grunt at the end of expiration. Ten dimensions pertain to *prosody*: rate, short phrases, increased rate in segments, increased rate overall, reduced stress, variable rate, prolonged intervals, inappropriate silences, short rushes of speech, and excess and equal stress. Five dimensions pertain to *articulation*: imprecise consonants, prolonged phonemes, repeated phonemes, irregular articulatory breakdowns, and distorted vowels. Finally, two general impression dimensions: *intelligibility* and *bizarreness*. The dimensions were then rated on severity across all dysarthrias.

To do this, Darley, Aronson and Brown (1969a) used a 7-point equal-appearing interval scale to rate severity, where one represented normal speech and seven represented very severe deviation from normal. The authors listened to a series of speech samples and rated each speaker on one dimension at a time.

To determine intrarater reliability, 30 patients were rated twice on each of the 38 dimensions by each author with no significant differences between ratings. For interrater reliability, the three listeners agreed on 84% of the sets that the sample was either normal or abnormal. On the degree of severity, the listeners marked the same scale value or were within one scale value on 84 % of the sets. Dimensions that had a mean scale value (mean of the scaled

severity assigned by the three listeners) of 2.0 and above were considered most deviant for each neurologic group. The most deviant speech dimensions of the etiologies are summarized in Table 1.

Clinically, the methods developed by DAB (henceforth the Mayo Clinic rating system) are routinely used to identify the types of dysarthria and assess and treat patients with dysarthria. This approach is favorable to physiologic and acoustic methods due to the relative ease and lack of required materials (Fawcett, 2010). However, to justify using this system in isolation, it must be proven effective and reliable, given that 1) listeners in the original DAB study were not blinded to the etiologies of the participants, 2) the reliability was estimated in a relatively less conservative manner, and 3) only three listeners (the authors of the study) were included.

Table 1. Most deviant dimensions per etiology and related dysarthria

Etiology	Most Deviant Dimensions in Descending Order	Related Dysarthria
Bulbar Palsy	Hypernasality, strain-strangled voice, imprecise consonants, breathy voice, and monopitch	Flaccid
Pseudobulbar Palsy	Imprecise consonants, monopitch, reduced stress, harsh voice, monoloudness, low pitch, slow rate, hypernasality, strained-strangled voice, and short phrases	Spastic
Amyotrophic Lateral Sclerosis	Imprecise consonants, hypernasality, harsh voice, slow rate, monopitch, short phrases, distorted vowels, low pitch, monoloudness, excess and equal stress, and prolonged intervals	Mixed (flaccid and spastic)
Cerebellar Disorder	Imprecise consonants, excess and equal stress, irregular articulatory breakdowns, distorted vowels, and harsh voice	Ataxic
Parkinsonism	Monopitch, reduced stress, monoloudness, imprecise consonants, inappropriate silences, short rushes of speech, harsh voice, and breathy voice (continuous)	Hypokinetic
Dystonia	Imprecise consonants, distorted vowels, harsh voice, irregular articulatory breakdowns, strain-strangled voice, monopitch, and monoloudness	Hyperkinetic
Chorea	Imprecise consonants, prolonged intervals, variable rate, monopitch, harsh voice, inappropriate silences, distorted vowels, and excess loudness variation	Hyperkinetic

Furthermore, as noted above, the dimensions are consistently used in the clinical setting in the differential diagnosis of dysarthria. However, not all 38 dimensions are used with each case of dysarthria due to the impracticality of applying such a large number of dimensions to every patient. Therefore, it would be beneficial to establish which dimensions have the greatest reliability scores and are most salient to each of the dysarthrias. This would allow for fewer dimensions to be analyzed when differentiating a dysarthric patient, making the system more practical for the clinical setting.

Aim of Study

The specific aim of this study was to assess the inter- and intra- rater reliability of a subjective, perceptual-auditory rating system to accurately identify the deviant speech characteristics associated with hypokinetic dysarthria in the Parkinson's disease population. Based on a review of the literature, it is hypothesized that inexperienced listeners will be able to identify the deviant speech dimensions of hypokinetic dysarthria identified by Darley et al. (1965a; 1965b), will have sufficient intrarater reliability among the listeners, and the most salient and reliable dimensions for the description of hypokinetic dysarthria will be derived from interrater reliability.

Research Questions

- Do the naïve listeners reliably identify the deviant speech dimensions of hypokinetic dysarthria?
- Which dimensions are relatively more salient and reliable than others for hypokinetic dysarthria?

LITERATURE REVIEW

Reliability of perceptual ratings of dysarthria

A review of the relatively small literature reflects conflicting results on the reliability of the perceptual nature of the Mayo Clinic rating system. Only two studies, Zyski and Weisiger (1987) and Zeplin and Kent (1996), attempted to directly replicate the work of Darly, Aronson, and Brown (1969a; 1969b) to establish interrater reliability. These studies used the original speech samples collected by Darly, Aronson, and Brown in 1969. However, the listeners did not have prior knowledge of the neurologic condition of the speakers as Darly, Aronson, and Brown did (Bunton, Duffy, Rosenbek, & Kent, 2007).

In the first study, Zyski and Weisiger (1987) did not include all 38 dimensions reported by Darley et al. (1969a). The authors reduced the dimensions to those that Darley et al. (1969a) reported as "more interesting", having a mean scale value of 2.0 or greater. This reduced the 38 dimensions to 16 (Darley et al., 1969a, p. 251), potentially excluding valuable dimensions for the detection of dysarthria. The authors believed these dimensions to be "more salient" and had "greater power to differentiate types of dysarthia" (Zyski & Weisiger, 1987, p. 369). Also, dimensions that occurred in more than four dysarthria types were not included. Features were rated on a 7-point scale (1= no deviance from normal, 7 = severe deviance from normal) for all forms of dysarthria. This study used three listener groups, two groups of experienced speech-language pathologists (Group 1 marked any dimension perceived to be present in each sample and Group 2 was instructed to pick a maximum of three dimensions present in each sample and to list either the dysarthria type or neurologic disease) and speech-language pathology graduate students (given five hours of training in perceptual analysis of dysarthria and the same task as the speech-language pathologist Group 2). The authors' purpose of having the listeners note deviant

speech characteristics was to "determine to what extent perceptual analysis alone could be used to differentiate specific types of dysarthria or neurologic disease" (Zyski & Weisiger, 1987, p. 373).

Based on the criteria created by DAB, Group 1 of the speech-language pathologists accurately identified 19% of dysarthria type, Group 2 of the speech-language pathologists accurately identified 55% of dysarthria type/disease, and Group 3, the speech-language pathology graduate students, accurately identified 56% of dysarthria type/disease. The authors did not report on the degree of reliability among the listeners (Zyski & Weisiger, 1987).

According to Bunton et al. (2007), "the authors' decision to focus their analysis on those features with the greatest variability likely contributed to lower correlations and the negative conclusion that the Mayo Clinic rating system was not sufficiently reliable for clinical purposes" (p. 1482).

In the second study, Zeplin and Kent (1996) also used the original speech samples collected by Darley et al. (1969a). Five participants, two students in their last year of a 2-year master's program in speech-language pathology and three speech-language pathology doctoral students, rated the speech samples on all of the original 38 dimensions, except for two, on a 7-point scale (1 = normal speech, 7 = very severe deviation from normal speech). The two dimensions not rated in this manner, loudness level and pitch, were rated on bipolar extremes (low and high pitch; soft and loud level), where the value 4 represented normal speech and the values of 1 and 7 represented deviations from normal. The authors found that listeners were able to identify key perceptual features of dysarthria and had good intrarater reliability, but there were significant differences for interrater reliability across the speech dimensions. In the results, dimensions with a standard deviation of 1.0 or less among the listener ratings were considered to

be the most reliable. Among the seven dysarthria types, spastic, ataxic, and hypokinetic types had more than 50% of listeners' ratings fall below one standard deviation, indicating high interlistener reliability. The remaining four dysarthria types however, had less than 50% of listeners' ratings fall below one standard deviation. In fact, mixed dysarthria had 19% of listeners' ratings fall greater than two standard deviations, indicating low reliability.

Two other studies by Kearns and Simmons (1988) and Shear, Adams, and Davis (1991) attempted to establish interrater reliability using the Mayo Clinic rating system focused on ataxic dysarthria alone and with the use of new dysarthric speech samples.

First, Kearns and Simmons (1988) used a perceptual characteristics protocol similar to the one created by Darley et al. (1969a), with the exception of rate and the addition of pitch variability, rapid rate, and slow rate. Experienced speech-language pathologists rated the dimensions. Results showed a mean overall reliability level of 82% between listeners, comparable to the 84% interlistener agreement reported by Darly et al. (1969a), and no differences in rater reliability across perceptual features (Bunton et al., 2007).

On the contrary, Shear et al. (1991) reported significant differences in rater reliability across the speech dimensions (Bunton et al., 2007). The authors instructed experienced speech-language pathologists to rate subjects with ataxic dysarthria on five speech dimensions: imprecise consonants, excess and equal stress, irregular articulatory breakdown, distorted vowels, and harsh voice. The authors chose these dimensions because they represent the most deviant speech characteristics associated with cerebellar lesions, the neurologic disorder related to ataxic dysarthria.

More recently, Bunton et al. (2007) investigated inter- and intra- rater agreement for the perceptual ratings of dysarthria using the Mayo Clinic rating system. The authors defined

interrater agreement as "the extent to which the different listeners tend to assign exactly the same rating to each object" (Bunton et al., 2007, p. 1482). In this study, two groups of listeners, inexperienced speech-language clinicians and experienced speech-language pathologists, rated 47 speakers with various types of dysarthria. The listener groups rated all 47 speakers on all 38 features presented by Darley et al. (1969a) using a 7-point scale (1 = normal, 7 = very severe deviation from normal), one dimension at a time. According to Bunton et al. (2007), results showed "reasonable levels of listener agreement for all 38 perceptual features, with no significant differences in rater agreement between listener groups or across individual perceptual features" (p. 1491). Individual perceptual features had a range of 32% to 100%, however only 11 features had more than 50% agreement among the listeners. This represents roughly 30% of perceptual features, which may not be sufficient evidence of acceptable levels of agreement.

Given the conflicting results these studies have produced, the reliability of the Mayo Clinic rating system has yet to be clearly established. The aim of this study was to further investigate the reliability of the Mayo Clinic rating system, focusing on one specific dysarthria, hypokinetic, most often associated with Parkinson's disease. Parkinson's disease was selected due to the accessibility of speech samples collected by the LSU Speech Acoustics Lab in association with a large study that examines cross-language characteristics of dysarthria secondary to Parkinson's disease (NIH-NIDCD 012405). College students with no or limited experience with dysarthric speech participated as listeners and these results could then be compared to similar, previous work.

Parkinson's disease, its speech, and speech-related findings

Parkinson's disease is a term usually used for parkinsonism of unknown cause that is responsive to the dopaminergic drug, levodopa. Parkinsonism is the more generic term used to

refer to conditions with different etiologies and pathophysiology than Parkinson's disease. These etiologies and pathophysiologies can include vascular conditions, Alzheimer's disease, toxic-metabolic conditions, trauma, infectious conditions, normal pressure hydrocephalus, and obstructive hydrocephalus (Duffy, 2005).

Parkinson's disease and parkinsonism are the most common causes of hypokinetic dysarthria. Parkinson's disease is a slowly progressive, idiopathic neurologic disease, characterized by hypokinesia (Darly, Aronson, & Brown, 1975). It affects about 1 to 2% of the population over the age of 50. Dysarthria is a late emerging sign of Parkinson's disease, but it affects about 90% of all cases over the course of the disease (Duffy, 2005). Parkinson's disease is characterized by the progressive loss of dopaminergic neurons, primarily in the substantia nigra pars compacta (Skodda, 2011). It is also characterized by nerve cell loss in the locus cerulus and a decrease of dopamine in the striatum (Duffy, 2005). The faces of patients with parkinsonism often appear mask-like with infrequent blinking (Darly, Aronson, & Brown, 1965b). In addition to symptoms such as muscular rigidity, tremor, bradykinesia, and postural instability, many parkinsonism patients will develop hypokinetic dysarthria (Skodda, 2011).

Hypokinetic dysarthria results from damage to certain parts of the extrapyramidal system. Damage to other parts results in hyperkinesia and hyperkinetic dysarthria. The extrapyramidal system consists of the basal ganglia, the paired substantia nigra, and subthalamic nuclei of the upper brain stem. This system regulates the muscle tone required for posture and changing position. It facilitates the freedom and automaticity of movements for skilled voluntary acts. Extrapyramidal disease results in a reduction of movements, called hypokinesia. Slowness of movement, limited range of motion, immobility, and paucity of movement, rigidity, loss of automatic movement, and a resting tremor, which is abated by movement, characterize

hypokinesia (Darley et al., 1965b; Darley et al., 1975). Of these, limited range of motion has the greatest affect on speech (Darley et al., 1975).

In their original study, Darley, Aronson, and Brown (1969a) found the most striking salient characteristics of hypokinetic dysarthria to be *monopitch, monoloudness, and reduced stress*. Monopitch and monoloudness severity were "decidedly greater" (p. 258) in parkinsonism than the neurologic groups previously reviewed by the authors. Other distinctive prosodic characteristics present were inappropriate silences, short rushes of speech, and variable rate. It is noted that hypokinetic dysarthria was the only dysarthria that was not characteristically slow, but typically, as a group, was rated as slightly fast. The authors also noted that imprecise consonants was a prominent characteristic, explaining that this apparently resulted from reduced "excursion of the articulators" (p. 258) rather than the rate of articulation. Harsh voice and breathy voice were also heard.

In a companion paper to Darley et al. (1969a), Darly et al., (1969b) used correlation matrices to demonstrate co-occurrence of deviant speech dimensions across the different types of dysarthrias. Eight distinctive clusters of dysfunctions were discovered. The cluster that emerged for parkinsonism included the following speech dimensions: monopitch, monoloudness, reduced stress, and short phrases. This cluster was expanded to add short rushes of speech, variable rate, and imprecise consonants. The authors attributed these dimensions to the reduced range of movement, rigidity of laryngeal musculature, and difficulty initiating movement common to Parkinson's disease.

METHODS

This study used a within group, reliability design to determine reliability coefficients for intrarater reliability and interrater reliability across the speech dimensions. Of the original 38 dimensions, 37 were selected for this study. The list of 37 dimensions was borrowed from Duffy (2005), which excluded *Bizarreness*, most likely due to its redundant nature. These speech dimensions represent the independent variables and the scores of deviant speech severity as rated by the listeners served as the dependent variable.

Speakers

Audio recordings of 11 participants were selected to conduct a retrospective analysis. As part of a larger study, speech materials were selected from the archived Louisiana State University (LSU) Motor Speech Database (NIH-NIDCD 012405, 2012-2016). The participants were classified into two groups. The first group consisted of 8 participants with a neurologic diagnosis of Parkinson's disease and a clinical diagnosis of hypokinetic dysarthria, four males and four females. These participants ranged in age from 45 to 85 (M= 67.3, SD= 12.3), while post-disease-onset times ranged from 2 to 24 years (M=9, SD=7.4). All Parkinson's disease participants reported language, hearing, and cognitive skills that were adequate for completing the task.

The second group consisted of three neurologically healthy control participants, two males and one female. These participants ranged in age from 62 to 64 (M= 62.7, SD= 1.2). No participants reported any history of speech, language, hearing, or cognitive deficits. Participant information is summarized in Appendix B.

Recording Procedures

Audio recordings were obtained individually in a single session in a quiet setting (in a quiet room in their homes or in a sound-treated booth). Speech samples were collected either with a Perception 120 (AKG) microphone directly transferred to a Dell OptiPlex 750 computer or a professional portable recording device (TASCAM DR-40). Speech recordings were made with a sampling rate of 22.1 kHz and 16-bit quantization. The participants were instructed to read the Caterpillar Passage (Patel et al., 2013) aloud. The Caterpillar Passage was designed to be balanced in length with breadth of tasks for clinical efficiency, have comprehensive phonotactic coverage, include word and sentence form that examine respiratory, phonatory, articulatory, resonatory, and prosodic control, have isolated speech motor tasks for comparison within connected speech, and use contemporary vocabulary and simple syntax to focus on speech production abilities while minimizing cognitive load (Patel et al., 2013).

Listeners and Procedures

One hundred forty-eight inexperienced, undergraduate students majoring in communication sciences and disorders at Louisiana State University volunteered to participate in this study and served as listeners. The listeners were blind to the neurologic diagnosis of the speakers. Due to the inexperience of the listeners, each dimension was explained one at time and listeners were given an opportunity to ask questions for any further clarifications. Along with the explanation, the listeners were given a typed key that included written definitions of each dimension taken from Darley et al. (1969a). Listening sessions ranged from individual sessions to large groups of up to 80 listeners in a quiet room. The audio recordings were played to each group of listeners from an audio file on the software TF32 (Milenkovic, 2005), inside a quite

listening setting. Each rating sessions lasted approximately 1 hour and the listeners were given the option for breaks as needed.

The listeners were given verbal instructions and told that they would hear the Caterpillar Passage read aloud by different speakers. For each speaker, the listeners were given a rating sheet containing an equal-appearing interval scale from 1 to 7. On this scale, 1 represented normal speech and 7 represented severe deviation from normal. The listeners were asked to use this scale to rate each speaker in all of the 37 dimensions. The listeners rated one speaker in every dimension before moving on to the next speaker. Listeners rated the following 37 dimensions: abnormal pitch, pitch breaks, monopitch, voice tremor, monoloudness, excess loudness variation, loudness decay, alternating loudness, loudness level (overall), harsh voice, hoarse (wet) voice, breathy voice (continuous), breathy voice (transient), strained (strained-strangled) voice, voice stoppages (interruptions/arrests), hypernasality, hyponasality, nasal emission, forced inspiration-expiration, audibly inspiration, grunt at the end of expiration, rate (slow or fast), short phrases, increased rate in segments (accelerated rate), increased rate overall (rapid rate), reduced stress, variable rate, prolonged intervals, inappropriate silences, short rushes of speech, excess and equal stress, imprecise consonants/articulation, prolonged phonemes, repeated phonemes or syllables, irregular articulatory breakdowns, distorted vowels, and speech intelligibility overall. These dimensions represent the 38 originally presented by Darley et al. (1969a), with the exception of *bizarreness*. See Appendix C for definitions of dimensions given to the listeners. Definitions were taken from Duffy (2005).

The same order of speakers was presented to each listener. The order of the speakers was randomly generated using RANDOM.ORG – List Randomizer, which included both groups of speakers, with and without dysarthria.

During rating sessions, the speakers' reading passages were presented once. Listeners were instructed to ask for the recording to be replayed if he/she needed to hear it again. Due to the number of dimensions being rated, listeners were allowed as much time as necessary to rate each dimension per audio sample. After each listener in the session had rated each dimension, the next sample was played.

To establish intrarater reliability, 2 speakers' passages were duplicated. One healthy speaker's passage and one dysarthic speaker's passage were selected from the middle of the randomized order and added to the end of the order. In total, the listeners rated 13 readings of the Caterpillar Passage. The duplication of 2 of the 11 speech samples represents more than the standard 5% typically used to determine intrarater reliability. Raw scores obtained for the ratings of the 37 dimensions were analyzed for intra- and interrater reliability.

Analysis

Reliability. The results of the 148 individual raters for each of the 37 dimensions per speaker (148 listeners x 37 dimensions x 13 speech samples = 71,188 ratings) were put into a spreadsheet for statistical analysis. The order of speech samples was randomly generated. The two duplicate speech samples were selected from the middle of the list, so that they would be least recognizable, and added to the end. The two duplicate samples represent about 20% of the speech samples. The listeners' ratings for the two sets of duplicates were statistically analyzed using Pearson product-moment correlation coefficient to determine intrarater reliability. The Pearson product-moment correlation coefficient is a measure of the strength of the linear relationship between two variables and is designated by *r* when measured in a sample. Pearson's *r* can range from -1 to 1, with -1 indicating a perfect negative linear relationship, 1 indicating a perfect positive relationship, and a value of 0 indicating no association between variables

(Kreinovich, Hung, & Berlin, 2013). A listener with a coefficient of at least 0.60 with a statistical significance level of $p < 0.05$ was considered to be sufficiently reliable for inclusion.

Interrater reliability was determined using intraclass correlation (ICC). The ICC coefficient is a measure of the reliability of ratings of two or more raters and describes how strongly units in the same group resemble each other. An ICC coefficient is measured on a 0 to 1 scale, where 1 indicates a perfect relationship and 0 indicates no relationship between units (Cleophas, Zwinderman, & Cleophas, 2002).

Saliency. Saliency was determined using similar methods as Darley et al. (1969a). DAB calculated the mean score value (MSV) by finding the average rating of the three judges for each of the original 38 dimensions. According to DAB, any MSV equal to or greater than 2.0 on the 7 point scale was considered a striking dimension. In this investigation, the MSV was calculated for both groups, Control and Parkinson's.

To further examine the difference in MSVs between the two speaker groups, an independent-samples t-test was conducted to compare the MSV of the PD speaker group and the MSV of the Healthy Control group for each dimension using SigmaPlot software.

RESULTS

The perceptual rating scores obtained from a total of 148 listeners on 37 dimensions were analyzed to determine (1) listener reliability (intra- and inter-) and (2) salient perceptual features of hypokinetic dysarthria secondary to Parkinson's disease.

Reliability

To determine intra-rater reliability, the two sets of duplicate speaker samples were analyzed using Pearson's coefficient for each listener. A listener with a coefficient of at least 0.60 with a statistical significance level of $p < 0.05$ was considered to be sufficiently reliable for inclusion (Anand & Stepp, 2015). Of the 148 listeners, 52 demonstrated reliability that met these standards.

Inter-rater reliability of the listeners was computed only for the 52 listeners who met the intra-rater reliability criteria by calculating the interclass correlation coefficient (ICC) for each of the 37 speech dimensions. The results are summarized in Table 2, which presents the dimensions in descending order of interrater reliability. The dimensions presented toward the top of the list with greater ICC coefficients are thought to be more reliable than dimensions toward the bottom of the table when applied for ratings of hypokinetic dysarthria.

Saliency

According to DAB (1969a), the most striking and salient features of a dysarthria were those that were determined to have a MSV of 2.0 or greater. In this investigation, perceptual saliency of Parkinson speech was identified in two ways. First, following the classic study by Darley et al. (1969a), the dimensions with the MSV of 2.0 or greater were considered "salient". Table 3 lists the dimension in descending order, beginning with the greatest MSV values. Second, the distance of the mean scores of the MSV between speakers with PD and healthy

speakers were calculated and a series of independent-sample t-test was conducted. Table 4 summarizes these results with the dimensions ranked by greatest *t* value in descending order. Figure 1 displays the MSV for each group per dimension with standard deviation.

Table 2. Speech dimensions ranked by correlation coefficient in descending order.

Rank	Speech Dimension	Corr. Coefficient
1	Harsh Voice	0.730
2	Speech Intelligibility Overall	0.606
3	Strained (strained-strangled) Voice	0.556
4	Voice Tremor	0.525
5	Monopitch	0.430
6	Short Rushes of Speech	0.392
7	Voice Stoppages	0.380
8	Imprecise Consonants/Articulation	0.370
9	Reduced Stress	0.357
10	Abnormal Pitch	0.347
11	Forced Inspiration-Expiration	0.315
12	Monoloudness	0.309
13	Irregular Articulatory Breakdowns	0.304
14	Increase Rate in Segments (accelerated rate)	0.300
15	Rate, slow or fast	0.299
16	Distorted Vowels	0.297
17	Short Phrases	0.286
18	Variable Rate	0.237
19	Increased Rate Overall (rapid rate)	0.230
20	Repeated Phonemes or Syllables	0.221
21	Hoarse (wet) Voice	0.219
22	Breathy Voice (continuous)	0.218
23	Audible Inspiration	0.216
24	Loudness Decay	0.192
25	Pitch Breaks	0.181
26	Excess and Equal Stress	0.163
27	Inappropriate Silences	0.155
28	Loudness Level (overall)	0.133
29	Prolonged Phonemes	0.123
30	Nasal Emissions	0.119
31	Prolonged Intervals	0.117
32	Grunt at End of Expiration	0.110
33	Alternating Loudness	0.098
34	Breathy Voice (transient)	0.095
35	Hyponasality	0.068
36	Excess Loudness Variation	0.068
37	Hypernasality	0.007

Table 3. Speech dimensions ranked in descending order by mean score value for Parkinson speakers. Dimensions with MSVs greater than 2.0 are considered striking or salient. The line between ranks 24 and 25 indicates the cutoff point for saliency per the DAB's criteria.

Rank	Speech Dimension	Mean Score Value
1	Monopitch	3.43
2	Strained (strained-strangled) Voice	3.2
3	Monoloudness	3.07
4	Harsh Voice	2.91
5	Rate, fast or slow	2.86
6	Reduced Stress	2.76
7	Voice Tremor	2.72
8	Imprecise Consonants/Articulation	2.68
9	Speech Intelligibility Overall	2.67
10	Voice Stoppages	2.57
11	Short Phrases	2.5
12	Short Rushes of Speech	2.49
13	Forced Inspiration-Expiration	2.42
14	Audible Inspiration	2.36
15	Hoarse (wet) Voice	2.31
16	Breathy Voice (continuous)	2.3
17	Abnormal Pitch	2.28
18	Variable Rate	2.28
19	Irregular Articulatory Breakdowns	2.28
20	Increased Rate in Segments	2.24
21	Excess and Equal Stress	2.23
22	Loudness Level (overall)	2.08
23	Pitch Breaks	2.06
24	Distorted Vowels	2.03
25	Loudness Decay	1.99
26	Breathy Voice (transient)	1.99
27	Increased Rate Overall (rapid rate)	1.91
28	Prolonged Intervals	1.78
29	Alternating Loudness	1.75
30	Repeated Phonemes or Syllables	1.72
31	Inappropriate Silences	1.71
32	Prolonged Phonemes	1.62
33	Hypernasality	1.56
34	Excess Loudness Variation	1.54
35	Grunt at the End of Expiration	1.46
36	Nasal Emission	1.38
37	Hyponasality	1.34

Table 4. Speech dimensions ranked by *t* value in descending order.

Rank	Speech Dimension	*t* Value
1	Speech Intelligibility Overall	$t(570) = 13.29, p < .001$
2	Strained (strained-strangled) Voice	$t(567) = 13.25, p < .001$
3	Monopitch	$t(569) = 12.18, p < .001$
4	Reduced Stress	$t(569) = 11.94, p < .001$
5	Rate, slow or fast	$t(570) = 10.57, p < .001$
6	Short phrases	$t(567) = 10.57, p < .001$
7	Imprecise consonants/articulation	$t(561) = 10.55, p < .001$
8	Harsh voice	$t(565) = 10.49, p < .001$
9	Voice stoppages	$t(566) = 10.42, p < .001$
10	Monoloudness	$t(569) = 9.42, p < .001$
11	Short rushes of speech	$t(568) = 9.41, p < .001$
12	Irregular articulatory breakdowns	$t(568) = 9.36, p < .001$
13	Forced inspiration-expiration	$t(567) = 8.91, p < .001$
14	Variable rate	$t(570) = 8.61, p < .001$
15	Hoarse (wet) voice	$t(568) = 8.09, p < .001$
16	Breathy voice (continuous)	$t(566) = 7.98, p < .001$
17	Distorted vowels	$t(569) = 7.56, p < .001$
18	Prolonged intervals	$t(568) = 7.45, p < .001$
19	Abnormal Pitch	$t(572) = 7.41, p < .001$
20	Excess and equal stress	$t(566) = 7.21, p < .001$
21	Pitch breaks	$t(568) = 6.98, p < .001$
22	Repeated phonemes or syllables	$t(570) = 6.78, p < .001$
23	Loudness decay	$t(570) = 6.70, p < .001$
24	Inappropriate silences	$t(569) = 6.51, p < .001$
25	Increased rate in segments	$t(569) = 6.14, p < .001$
26	Increased rate overall (rapid rate)	$t(568) = 5.82, p < .001$
27	Breathy voice (transient)	$t(565) = 5.75, p < .001$
28	Prolonged phonemes	$t(566) = 5.63, p < .001$
29	Audible inspiration	$t(570) = 5.23, p < .001$
30	Grunt at end of expiration	$t(565) = 5.14, p < .001$
31	Loudness level (overall)	$t(568) = 4.98, p < .001$
32	Voice tremor	$t(310) = 4.57, p < .001$
33	Nasal emission	$t(569) = 4.05, p < .001$
34	Hyponasality	$t(568) = 3.67, p < .001$
35	Hypernasality	$t(568) = 2.80, p < .05$
36	Excess loudness variation	$t(569) = 1.88, p = 0.0605$
37	Alternating loudness	$t(570) = 1.36, p = 0.105$

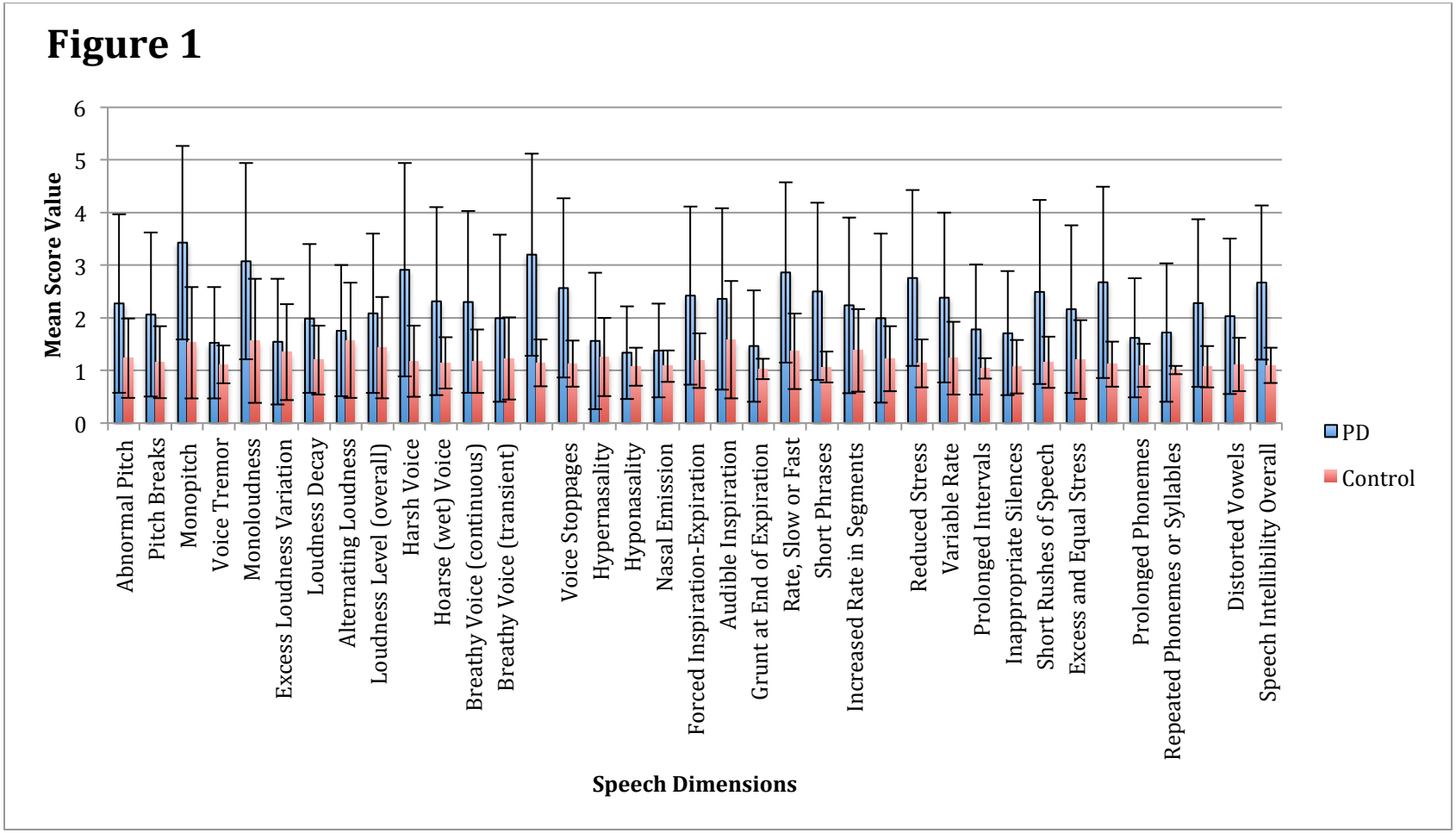

Figure 1. Mean Values of Parkinson's disease group compared to Healthy Control group per dimension.

DISCUSSION

This study sought to investigate and answer two questions: 1) Do the naïve listeners reliably identify the deviant speech dimensions of hypokinetic dysarthria? 2) Which dimensions are relatively more salient and reliable than others for hypokinetic dysarthria?

Analysis of interrater reliability and saliency for hypokinetic dysarthria suggested a number of speech dimensions that can be interpreted with greater reliability and saliency than others when describing hypokinetic dysarthria.

Reliability

Overall, one-third of the listener participants demonstrated acceptable intrarater reliability when the "0.6" or greater criteria was employed across all 37 dimensions. There are factors that could possibly affect this finding such as the listeners' experience with dysarthria, pool of listeners, and duration of experiment, although there is no agreement yet how these factors affect the reliability. For instance, it is not well understood how listeners' experience with dysarthria affect the reliability of perceptual ratings of dysarthria. Bunton et al. (2007) reported no significant difference in rater agreement between an experienced rater group and an inexperienced rater group. Conversely, when examining the role of experience in perception of phonetic detail, Munson, Johnson, and Edwards (2012) found that experienced SLPs demonstrated higher intrarater reliability than inexperienced listeners.

The degree of interrater reliability of a clinically applicable rating scale is important, especially in that these ratings serve as a basis of future treatment and management of treatment plan. A high level of reliability indicates that listeners tend to assign closely related meaning to each point of a scale, i.e. *normal, extremely abnormal,* and each point in between have similar meanings to each listener (Bunton et al., 2007). Four dimensions had a reliability coefficient greater than 0.500. This number of dimensions is comparable to the 11 dimensions found by

Bunton et al. (2007) with listener agreement greater than 50%. However, the findings of both this study and that of Bunton et al. (2007) suggest fewer dimensions with relatively high reliability than what has been reported in previous studies (Darley et al., 1969a; Kearns & Simmons, 1988; Sheard et al., 1991). However, there is not an established rule for what constitutes sufficient interrater reliability or agreement, such as a coefficient of 0.70 or agreement of 70%. Therefore, a sufficient level of reliability may differ among studies and authors. As such, the results of this investigation produced a continuum of reliability for the dimensions that can be interpreted by the user in selection of dimensions to be applied when rating hypokinetic dysarthric speech.

Interestingly, greater reliability was not always found at the extreme points of the scale (i.e., *1=normal*, *7 = extremely abnormal*). For example, *hyponasality* was determined to have a MSV of 1.34 for the Parkinson speaker group, which was the closest MSV to normal of all 37 dimensions (Healthy speakers: 1.07). These results would indicate most listeners determined that the Parkinson group spoke with normal nasality or that it was not deviantly hyponasal, which would then lead one to expect that this dimension demonstrated higher reliability across the listeners. However, *Hyponasality* was determined to have one of the lowest scores for interrater reliability (ICC = 0.068). The results for this dimension contradict the expectation of higher reliability at the endpoints of the scale, indicating some listeners may not have clearly understood this feature or may have confused the representation of the scale (e.g., 1= severe deviation, 7= normal) which would produce extreme outliers that could have skewed the analysis.

Given the results of this investigation, when taking interrater reliability into consideration for the clinical use of the 37 speech dimensions for the rating of hypokinetic dysarthria it is recommended to use *Harsh Voice*, *Speech Intelligibility Overall*, *Strained (strained-strangled Voice*, and *Voice Tremor*. These dimensions were found to have the highest listener reliability.

Salient Features of Hypokinetic Dysarthria

According to DAB, the most salient and striking characteristics of hypokinetic dysarthria, which showed a mean score value of 2.0 or greater, included *Monopitch*, *Monoloudness*, *Reduced Stress*, *Imprecise Consonants/Articulation*, *Inappropriate Silences*, *Short Rushes of Speech*, *Harsh Voice*, and *Breathy Voice (continuous)*. The results of the current study identified 24 dimensions as having a mean score value of 2.0 or greater for the Parkinson group. Of the 9 dimensions identified by DAB as salient for hypokinetic dysarthria, all but *Inappropriate Silences* were identified by the listeners in the current study as being salient. The overall number of dimensions identified by the listeners can be interpreted as the listeners being able to differentiate unaffected speech from affected speech. The difference in the number of salient features between this investigation and that of DAB could indicate the listeners had difficulties determining the more significantly affected dimensions of speech; however, the listeners' ratings could have been influenced by severity if the speakers in this study were more severe than the ones in DAB's study. Therefore, a more detailed investigation of mean score values was needed. Table 5 summarizes the most salient features identified by DAB in comparison to the top 10 most salient features identified in this study.

Unlike DAB, the present investigation included a healthy control group of speakers that were rated on the 37 dimensions as well. For a more objective analysis of the mean score values, the means for each group per dimension were compared using an independent-sample t-test. The results indicated that the MSVs of the Parkinson group were statistically different from the MSVs of the Healthy Control group for 35 dimensions ($n = 34$ with $p < 0.001$; $n = 1$ with $p < 0.05$). Two dimensions, *excess loudness variation* and *alternating loudness*, did not have a statistical difference between the two group's MSVs. These two dimensions also had MSVs below 2.0.

Table 5. Comparison of most salient features identified by DAB and top 10 most salient identified in present study.

DAB Most Salient Features	Current Study Top 10 Most Salient Features
*Monopitch**	*Monopitch**
*Monoloudness**	Strained (strained-strangled) Voice
*Reduced Stress**	*Monoloudness**
*Imprecise Consonants/Articulation**	*Harsh Voice**
Inappropriate Silences	Rate, fast or slow
Short Rushes of Speech	*Reduced Stress**
*Harsh Voice**	Voice Tremor
Breathy Voice (continuous)	*Imprecise Consonants/Articulation**
	Speech Intelligibility Overall
	Voice Stoppages

*****Identified as salient in both studies.

The results of the t-test further reinforce which dimensions are most salient to hypokinetic dysarthria. The dimensions with the greatest mean score values that also have the greatest *t* values can be concluded to be the most salient deviant dimensions. For example, *monopitch* was the dimension with the greatest MSV and also has the third greatest *t* value, indicating this dimension was subjectively and objectively one of the most deviant speech dimensions identified by the listeners for hypokinetic dysarthria. In addition, the overall similarity between the two studies with respect to most salient speech characteristics of Parkinson's disease (including speech dimensions and their ranks) concurs that naïve listeners are able to identify prominent speech disturbances of dysarthria (at least related to Parkinson's disease).

Given the results of this investigation, when taking saliency into consideration for the use of the 37 speech dimensions in rating hypokinetic speech the use of *Monopitch*, *Strained (strained-strangled) Voice*, *Monoloudness*, *Harsh Voice*, and *Rate, fast or slow* is recommended as these have the highest MSVs that were statistically different from the control group.

Clinical Implications

The results of this investigation are of great consequence in the clinical setting. The Mayo Clinic rating system is considered by many to be a gold standard in the classification of dysarthria type; however it is not often implemented due to the large scale of dimensions to be rated, making it less practical for everyday use. The results of this investigation, in conjunction with future research regarding the remaining dysarthria types, may be able to reduce the number of dimensions to the most reliable and salient of each dysarthria to produce a more practical tool for the clinical setting. This tool could potentially make identifying the patterns of deviant speech for each dysarthria type more efficient.

In consideration of reliable and efficient assessment of speech characteristics associated with Parkinson's disease, the dimensions *Harsh Voice*, *Strained (strained-strangled) Voice*, *Monopitch*, *Monoloudness*, *Rate, fast or slow, Speech Intelligibility Overall, and Voice Tremor* are recommended following the current results.

Limitations

The hypokinetic dysarthic speakers who participated in this study ranged from mild to moderate in severity. This was due to the severe speaker's sample that was available being so severe that it did not meet the parameters of the study. A more severe speech sample may have given the listeners a better reference point for the severe deviation from normal speech aspect of the rating scale. That is to say, with the presentation of a severe speech sample the listeners may have rated the mild to moderate sample less harshly, potentially reducing the mean score values for some speech dimensions. Therefore, there would be more differentiation between the non-striking and salient speech dimensions of hypokinetic dysarthria.

Future Research

The findings from this investigation provide support of an auditory-perceptual rating scale for evaluation of hypokinetic dysarthic speech; however, there is question regarding the reliability of this system. Future research should investigate the most reliable and salient speech dimensions for the other dysarthria types. This would ideally allow for a fewer number of dimensions to be analyzed to classify the dysarthrias, making it a more efficient and practical tool to be used in the clinical setting. Finally, cluster analysis of the 37 dimensions would be another approach to developing a more efficient tool for perceptual evaluation, which is considered as its next step of this study.

\

References:

Anand, S., & Stepp, C. E. (2015). Listener Perception of Monopitch, Naturalness, and Intelligibility for Speakers With Parkinson's Disease. *J Speech Lang Hear Res Journal of Speech Language and Hearing Research, 58*(4), 1134.

Bunton, K., Kent, R. D., Duffy, J. R., Rosenbek, J. C., & Kent, J. F. (2007). Listener Agreement for Auditory-Perceptual Ratings of Dysarthria. *Journal of Speech, Language & Hearing Research*, *50*(6), 1481-1495. doi:10.1044/1092-4388(2007/102).

Cleophas, T. J., Zwinderman, A. H., & Cleophas, T. F. (2002). *Statistics applied to clinical trials*. Dordrecht: Kluwer Academic.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12*, 246-269.

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

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). Hypokinetic Dysarthria: Disorders of the Extrapyramidal System. In *Motor Speech Disorders* (pp. 171-197). Philadelphia, Pennsylvania: Saunders.

Dorland, W. (1965). Dorland's Illustrated Medical Dictionary (24th ed.). Philadelphia & London: W. B. Saunders.

Duffy, J. (2005). Hypokinetic Dysarthria. In *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, Missouri: Elsevier Mosby.

Kearns, K., & Simmons, N. (1988). Interobserver Reliability and Perceptual Ratings: More Than Meets the Ear. *Journal of Speech and Hearing Research, 31*, 131-136.

Kreinovich, V., Hung T., N., & Berlin, W. (2013). Towards a Localized Version of Pearson's Correlation Coefficient. *International Journal of Intelligent Technologies & Applied Statistics*, *6*(3), 215-224.

Milenkovic, P. (2005). TF32 [computer software]. Madison, WI: University of Wisconsin-Madison.

Munson, B., Johnson, J. M., & Edwards, J. (2012). The Role of Experience in the Perception of Phonetic Detail in Children's Speech: A Comparison Between Speech-Language Pathologists and Clinically Untrained Listeners. *American Journal of Speech-Language Pathology*, *21*(2), 124-139.

Patel, R., Connaghan, K., Franco, D., Edsall, E., Forgit, D., Olsen, L., . . . Russell, S. (2013). "The Caterpillar": A Novel Reading Passage for Assessment of Motor Speech Disorders. *American Journal of Speech-Language Pathology, 22*, 1-9.

RANDOM.ORG – List Randomizer. Retrieved January 19, from 2015,https://www.random.org/lists/.

Sheard, C., Adams, R., & Davis, P. (1991). Reliability and Agreement of Ratings of Ataxic Dysarthric Speech Samples with Varying Intelligibility. *Journal of Speech and Hearing Research, 34*, 285-293.

Skodda, S. (2011). Aspects of speech rate and regularity in Parkinson's disease. *Journal of the Neurological Sciences*, 310(1/2), 231-236.

Zaiontz, C. (2014). Intraclass Correlation. Retrieved May 2, 2015, from http://www.real-statistics.com/reliability/intraclass-correlation/.

Zeplin, J. & Kent, R. D. (1996). Reliability of auditory-perceptual scaling of dysarthria. In D.A. Robin, K.Y. Yorkston, & D. R. Beukelman (Ed.), Disorders of motor speech. Paul H. Brooks Publishing Co.: Baltimore.

Zyski, B. J., & Weisiger, B. E. (1987). Identification of dysarthria types based on perceptual analysis. *Journal of Communication Disorders, 20*, 367-378.

APPENDIX A. DIMENSIONS USED IN DAB ORIGINAL STUDY

No.	Dimension	Description
1.	Pitch level	Pitch of voice sounds consistently too low or too high for individuals age and sex.
2.	Pitch breaks	Pitch of voice shows sudden and uncontrolled variation (falsetto breaks).
3.	Monopitch	Voice is characterized by a monopitch or monotone. Voice lacks normal pitch and inflectional changes. It tends to stay at one pitch level.
4.	Voice tremor	Voice shows shakiness or tremulousness.
5.	Monoloudness	Voice shows monotony of loudness. It lacks normal variations in loudness.
6.	Excess loudness variation	Voice shows sudden, uncontrolled alterations in loudness, sometimes becoming too loud, sometimes too weak.
7.	Loudness decay	There is progressive diminution or decay of loudness.
8.	Alternating loudness	There are alternating changes in loudness.
9.	Loudness (overall)	Voice is insufficiently or excessively loud.
10.	Harsh voice	Voice is harsh, rough, and raspy.
11.	Hoarse (wet) voice	Wet, “liquid sounding” hoarseness.
12.	Breathy voice (continuous)	Continuously breathy, weak, and thin.
13.	Breathy voice (transient)	Breathiness is transient, periodic, and intermittent.
14.	Strained-strangled voice	Voice (phonation) sounds strained or strangled Ian apparently effortful squeezing of voice through glottis).
15.	Voice stoppages	There are sudden stoppages of voiced air stream (as if some obstacle along vocal tract momentarily impedes flow of air).
16.	Hypernasality	Voice sounds excessively nasal. Excessive amount of air is resonated by nasal cavities.
17.	Hyponasality	Voice is denasal.
18.	Nasal emission	There is nasal emission of air stream.
19.	Forced inspiration-expiration	Speech is interrupted by sudden, forced inspiration and expiration sighs.
20.	Audible inspiration	Audible, breathy inspiration.
21.	Grunt at end of expiration	Grunt at end of expiration.
22.	Rate	Rate of actual speech is abnormally slow or rapid.
23.	Phrases short	Phrases are short (possibly due to fact that inspirations occur more often that normal). Speaker may sound as if he has run out of air. He may produce a gasp at the end of a phrase.
24.	Increase of rate in segments	Rate increases progressively within given segments of connected speech.
25.	Increase of rate overall	Rate increases progressively from beginning to end of sample.
26.	Reduced stress	Speech shows reduction of proper stress or emphasis patterns.
27.	Variable rate	Rate alternately changes from slow to fast.
28.	Intervals prolonged	Prolongation of interword or intersyllable intervals.
29.	Inappropriate silences	There are inappropriate silent intervals.
30.	Short rushes of speech	There are short rushes of speech separated by pauses.
31.	Excess and equal stress	Excess stress on usually unstressed parts of speech, e.g. (1) monosyllabic words and (2) unstressed syllables of polysyllabic words.
32.	Imprecise consonants	Consonant sounds lack precision. They show slurring, inadequate sharpness, distortions, and lack of crispness. There is clumsiness in

		going from one consonant sound to another.
33.	Phonemes prolonged	There are prolongations of phonemes.
34.	Phonemes repeated	There are repetitions of phonemes.
35.	Irregular articulatory breakdown	Intermittent nonsystematic breakdown in accuracy of articulation.
36.	Vowels distorted	Vowel sounds are distorted throughout their total duration.
37.	Intelligibility (overall)	Rating of overall intelligibility or understandability of speech.
38.	Bizarreness (overall)	Rating of degree to which overall speech calls attention to itself because of its unusual, peculiar, or bizarre characteristics.

APPENDIX B. SPEAKER INFORMATION

Speaker	Gender	Age	Onset of PD
PD1	M	55	24 years
PD2	M	85	5 years
PD3	M	74	8 years
PD4	M	74	4 years
PD5	F	69	13 years
PD6	F	68	2 years
PD7	F	68	3 years
PD8	F	45	13 years
HC1	F	64	---
HC2	F	62	---
HC3	M	62	---

APPENDIX C. DIMENSION DEFINITIONS GIVEN TO LISTENERS

1. **Abnormal pitch:** Pitch is consistently too low or high for age and sex.
2. **Pitch breaks:** Pitch shows sudden and uncontrolled variation (falsetto breaks).
3. **Monopitch:** Voice is characterized by monopitch or monotone. Voice lacks normal pitch variation.
4. **Voice tremor:** Voice shows fairly regular shakiness or tremor.
5. **Monoloudness:** Voice shows monotony of loudness. It lacks normal variations in loudness.
6. **Excess loudness variation:** Voice shows sudden, uncontrolled alterations in loudness, sometimes becoming too loud, sometimes too quiet.
7. **Loudness decay:** Progressive diminution or decay of loudness within an utterance.
8. **Alternating loudness:** Alternating changes in loudness within an utterance.
9. **Loudness level (overall):** Voice is insufficiently or excessively loud
10. **Harsh voice:** Voice is harsh, rough, and raspy.
11. **Hoarse (wet) voice:** There is wet, "liquid-sounding" hoarseness.
12. **Breathy voice (continuous):** Voice is continuously breathy, weak, and thin.
13. **Breathy voice (transient):** Breathiness is transient, periodic, and intermittent.
14. **Strained (stained-strangled) voice:** Voice quality sounds strained or strangled (an apparently effortful squeezing of voice through glottis).
15. **Voice stoppages (interruptions/arrests):** There are sudden stoppages of voice, as if airflow has been impeded.
16. **Hypernasality:** Resonance is excessively nasal.
17. **Hyponasality:** Resonance is hyponasal/denasal.
18. **Nasal emission:** There is nasal emission of air during speech, sometimes audible.
19. **Forced inspiration-expiration:** Speech is interrupted by sudden inspiration or expiration.
20. **Audible inspiration:** Audible, breathy inspiration.
21. **Grunt at end of expiration:** there is a grunt at the end of expiration during speech.
22. **Rate, slow or fast:** Rate of speech is abnormally slow or rapid.
23. **Short phrases:** Phrases are short (possibly because inspirations occur more often than normal). Speaker may sound as if he or she has run out of air.
24. **Increased rate in segments (accelerated rate):** Rate increases progressively within given segments of connected speech.
25. **Increased rate overall (rapid rate):** Rate increases progressively from beginning to end of sample.
26. **Reduced stress:** Speech shows reduction of proper stress or emphasis patterns.
27. **Variable rate:** Rate varies within or across utterances.

28. **Prolonged intervals:** There is prolongation of inter-word or inter-syllable intervals.
29. **Inappropriate silences:** There are inappropriate silent intervals.
30. **Short rushes of speech:** There are short, rapid rushes of speech separated by pauses.
31. **Excess and equal stress:** There is excess stress on usually unstressed syllables of parts of speech (e.g., unstressed syllables of polysyllabic words).
32. **Imprecise consonants/articulation:** Consonants lack precision. They show inadequate sharpness, distortions, and lack crispness.
33. **Prolonged phonemes:** Phonemes are prolonged.
34. **Repeated phonemes or syllables:** There are slow or rapid repetitions of phonemes.
35. **Irregular articulatory breakdowns:** There are intermittent, nonsystematic breakdowns in precisions of articulation.
36. **Distorted vowels:** Vowels are distorted in their phonetic accuracy.
37. **Speech intelligibility overall:** How well the speaker is understood.

VITA

Jessica Lynn Miller first attended Louisiana State University where she earned her Bachelor of Arts degree in Communication Disorders in December 2013. She began her Master of Arts degree in August 2014 at Louisiana State University and she is a candidate to graduate in May 2016. Her thesis was completed under the guidance of Dr. Yunjung Kim. Upon graduation, Jessica plans to work as a clinical fellow speech-language pathologist in a medical setting with a special interest in treating dysphagia patients.