7-23-2019 2:00 PM

Changes in speech intelligibility and acoustic distinctiveness along a speech rate continuum in Parkinson's disease

Thea Knowles
The University of Western Ontario

Supervisor
Adams, Scott G.
The University of Western Ontario

Graduate Program in Health and Rehabilitation Sciences
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Thesis abstract

Asking a person to speak slowly is a common technique in speech therapy for people with Parkinson's disease (PD). Slowed speaking rates are thought to bring about changes in speech production that make it easier for people with speech impairments associated with PD to be understood, but this is not always the case. Furthermore, research suggests that using faster speech does not necessarily lead to decreases in speech intelligibility for some people with PD. Most studies of rate modification in PD have only included one or two rate adjustments to investigate the relationship between speech rate, intelligibility, and acoustic aspects of speech production. The present study adds to this literature and expands it by eliciting a broader range of speech rates than has previously been studied in order to provide a comprehensive description of changes along such a continuum.

Two groups of people with PD and documented speech changes participated: 22 receiving standard pharmaceutical intervention, and 12 who additionally had undergone deep brain stimulation surgery (DBS), a common surgical treatment for PD. DBS is often associated with further speech impairment, but it is unknown to what extent these individuals may benefit from speech rate adjustments. Younger and older healthy control groups were also included. All participants were asked to modify their speech rate along a seven-step continuum from very slow to very fast while reading words, sentences, and responding to prompts. Naïve listeners later heard these speech samples and were asked to either transcribe or rate what they heard.

Results indicated different patterns of speech changes across groups, rates, and tasks. Sentence reading and conversational speech were rated as being more intelligible at slow rates, and less intelligible at fast rates. All modified rates were found to *negatively* impact speech sound identification during a novel carrier phrase task. Slower speech was overall associated with greater acoustic contrast and variability, lower intensity, and higher voice quality. Differences in acoustic speech adjustments across the groups and speech rates emerged, however, in particular for the DBS group. Findings pointed to a complex relationship between speech rate modifications, acoustic distinctiveness, and intelligibility.

Lay summary

Parkinson's disease (PD) is a neurodegenerative disorder that often is associated with changes to a person's speech. This makes it difficult for some people with PD to be understood when using speech to communicate. Speech-language pathologists who treat people with PD will often work with them to try and slow down their rate of speech. Slow speech is a common form of intervention that has been shown to improve spoken communication in people with PD, making it easier for them to be understood. However, not all people with PD benefit from using slow speech. Furthermore, speaking more quickly is not necessarily associated with speech that is more difficult to understand. The goal of this thesis was to explore speech changes that occurred in people with and without PD across many different speech rates from very slow to very fast in order to better understand these patterns.

Two groups of people with PD participated: 22 receiving standard antiparkinsonian medication, and 12 who additionally had undergone deep brain stimulation surgery (DBS), a common surgical treatment for PD. DBS is often associated with greater and more variable speech impairment. Younger and older healthy control groups were also included. All participants completed various speech tasks (i.e., sentence reading, nonsense word reading, and conversation) at seven different rates from very slow to very fast. Naïve listeners later heard these speech samples and were asked to either transcribe or rate what they heard.

Results indicated different patterns of speech changes across groups, rates, and tasks. Sentence reading and conversational speech were rated as being more understandable at slow rates, and less understandable at fast rates. Nonsense words were more difficult to understand at both slower and faster rates of speech compared to normal rates. Slower speech overall was produced more quietly, with greater hoarseness, and with more speech sound contrast compared to fast speech, though these patterns differed across groups. The findings suggest complex relationships between speech rate, speech characteristics, and understandability across the groups.

Keywords

Parkinson's disease, speech acoustics, speech rate, speech intelligibility, deep brain stimulation, mixed effects regression, aging

Dedication

For Leland.

Acknowledgments

First and foremost, thank you to my supervisor, Dr. Scott Adams, for your guidance, enthusiasm, support, and encouragement. You enthusiastically fostered my creativity and intellectual curiosity every step of the way. You make it easy to learn deeply and still be able to step back and see the bigger picture, to think critically and clinically. Thank you for your positivity and confidence in my abilities, for your constructive feedback and generous praise. You are my gold standard for what it is to be a mentor, and I will do my best to emulate the environment you have created for your students as I move on to the next stages of my career.

To the Parkinson's community in London and beyond, thank you for the opportunities to learn from you. To all those who participated, thank you for your time and dedication. Thank you for thought-provoking conversations and insights into what it is to be living with Parkinson's. I am a better researcher, clinician, and person thanks to these interactions. I am grateful to the Parkinson Society of Southwestern Ontario for providing funding opportunities for graduate students like myself.

My clinical research would not exist without Dr. Mandar Jog and the Movement Disorders Centre. Thank you for challenging all of us to think broadly and critically about movement disorders. Your dedication, passion, and expertise are invigorating and inspirational.

Dr. Allyson Page, I am so grateful for your mentorship, kindness, and encouragement. You ignite empathy in your students and have been a beacon of positivity for so many of us. Thank you for your thoughtful commentary and for always connecting the research to the important clinical questions. Thank you especially for being a person to turn to, especially this year; our conversations always leave me with a sense of calm and encouragement.

I am grateful to Dr. Vijay Parsa, for your thoughtful comments and questions while I developed the ideas in this thesis. Thank you to Dr. Yun-Hee Choi for support with the analysis and being so generous with your time and encouragement. Dr. Aditya Muragi, and Dr. Soumya Sharma: I always look forward to our discussions about clinical questions and implications and to watching and learning from you. Thank you to Greydon Gilmore for our stimulating discussions about movement disorders to statistical modelling to coffee.

Dr. Meghan Clayards, you were the spark that got me in to the world of speech acoustics. Thank you for providing me with those first formative opportunities working in your lab, learning Praat, and R, and how to be a researcher in the first place, as well as for all your guidance and support since. Dr. Morgan Sonderegger, thank you for driving me to push my skills beyond what I thought I could do. For always providing thoughtful, accessible explanations to difficult topics. Thank you to both of you for being an incredible team to learn from, prior to my PhD and all the way through it. I look forward to continued collaborations as I move on to the next stage.

I am incredibly grateful to Dr. Ken McRae and Dr. Angie Roberts for providing me with the opportunity to think about communication in Parkinson's disease through a lens I may not otherwise have had. Thank you for your high standards and supportive advice.

Thank you to the HRS and CSD faculty and staff. To Barb Nikolakakos; you are very often the best part of my day when I go to Elborn (even if I'm already having a great day). Thank you for your attention to detail, your compassion, your wit. Thank you for being you. You have brightened the days, weeks, months, lives of so many HRS students.

I am so grateful to all of the SLP students who participated and assisted on the project. Thank you for your willingness and interest during what I am all too aware is a very demanding and stressful couple of years.

What luck to have stumbled into the R-Ladies Global community. You all have created such a wonderful world to be a part of, to learn in, and to be vulnerable in. Thank you to the R user community in general for open source tools and such a supportive environment. A special thanks to our own R-Ladies #LdnOnt members for showing up and engaging, for being there to learn with and from.

To my lab mates and friends Anita, Cynthia, and Daryn: thank you for always being there to chat about speech or life, for being such wonderful pillars of support throughout these last few years. Further thanks to Daryn for your careful eye in looking over portions of this thesis and for many thoughtful discussions as well as much needed revels and rants.

This process would be a much lonelier one without my combined clinical program comrades as we weather the storm together. In particular to Sally, for leaping at my crazy idea to start an R-Ladies chapter. I could not have dreamed up a better partner in crime and code, and I am grateful for your friendship. To Jonny, for always being a person with whom I can commiserate but also celebrate accomplishments; thank you for running alongside me, quite literally, throughout these last few years. To Blair, for your kindness and compassion alongside your intellect and enthusiasm.
To my clinical cohort pals and winebookclub: Kelsey (& Darcy), Emily, Taylor, Scott, Chloe, Siobhan. How would I have made it through the first two years without you? Grateful you are all in my lives.

To Carol for always pushing me to be my best, for writing retreats and late-night conversations over the last decade. To Jennie and Mary, for your loving check-ins and for being your beautiful, thoughtful, empathetic selves. To my NAC team, years and miles away but always there to snark with, to reminisce, and to ride into every stage of life with. Pete, Frodge, K-Rex, Scammy, Leedog. Kitty, Wyatt, John: you are my village. Leland, you are my every day. I'm taking good care of your boy Rocky, but he's taking better care of me.

To my housemates, Brandon, Joe, Jordan, and Rachel, for making a home here with me in our own lost-toys kind of way. For making it easy to just be. Thank you for feeding Rocky cheese and always saying yes to burritos. Special thanks to Brandon for ensuring I was eating and sleeping during the final months of thesis writing.

To my parents and G-Love for being just the most supportive, encouraging, loving family I could have ever hoped to have been born into. Thank you for asking how my health and happiness are, and always being easy to talk to regardless of the answers to those questions. To my extended Leboldian and Knowlesian clans for your love, humour, and support.

Brian, for being my rock through it all, no matter where in the world we happen to land, no matter what in the world we have to face. Thank you for believing in me and in us.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

ACI: Acoustic correlates of intelligibility

AIC: Akaike information criterion

DAF: Delayed auditory feedback

dB: Decibel

DBS: Deep brain stimulation

F1: First formant

F2: Second formant

HkD: Hypokinetic dysarthria

HNR: Harmonics-to-noise ratio

ICC: Intraclass correlation coefficient

LSVT: Lee Silverman Voice Treatment

MFA: Montreal Forced Aligner

MoCA: Montreal Cognitive Assessment

OC: Older healthy controls

PD: Parkinson's disease

PoA: Place of articulation

QVAI: Quadrilateral vowel articulation index

SIT: Sentence Intelligibility Test

SLM: Sound level meter

SNR: Signal-to-noise ratio

UPDRS: Unified Parkinson's Disease Rating Scale

VAS: Visual analog scale

VDC: Voicing during closure

VOT: Voice onset time

WPM: Words per minute

YC: Younger healthy controls

1 Thesis overview

1.1 Objective

Speech rate modification is frequently used in behavioural speech therapy to improve the communication of people with Parkinson's disease (PD), but our understanding of the ways in which a person's speech changes along a continuum from very slow to very fast is not well understood. The overarching goal of this series of investigations is to describe acoustic and perceptual changes that occur along a range of possible speech rates, and how these changes differ for individuals with PD. Understanding speech modifications that occur along such a continuum will inform clinicians' approaches to treatment selection and determining candidacy for rate modification. This objective is accomplished in the present study by eliciting a continuum of speech rates, from very slow to very fast, in people with PD and hypokinetic dysarthria (HkD) as well as in control speakers. Two sub-groups of people with PD are of interest: people with PD and HkD who are undergoing standard pharmaceutical treatment, as well as people with PD and HkD who have undergone a common surgery known as deep brain stimulation (DBS) to treat the primary motor symptoms of PD. Individuals with PD who have received DBS often report worsening of speech symptoms over time. The primary group of control speakers is a cohort of age-matched healthy older adults. Younger adults are also included for a subset of the analyses in order to explore the effects of aging on speech modifications that occur as a function of speech rate.

1.2 Organization of dissertation

Chapter 2 discusses the nature of PD and the speech symptoms that are associated with the disease itself, as well as following DBS. Section 2.3 presents a review of the literature that has addressed speech rate modification as a therapeutic strategy, and discusses these findings in the context of rate-induced changes in speech acoustics (Section 2.3.3) and intelligibility (Section 2.3.2) both for individuals with dysarthria as well as neurologically healthy talkers.

Chapter 3 describes the methodology. The overall study is reported as three distinct experiments:

1. **Experiment 1**: Speech production: A systematic exploration of acoustic changes of speech production induced by speech rate modifications along a continuum from very slow to very fast.

2. **Experiment 2**: Speech transcription: A perceptual study in which listeners were asked to transcribe nonsense words produced in a subset of the tasks in Experiment 1. Speech intelligibility in terms of consonant and vowel accuracy are the outcomes of interest.

3. **Experiment 3**: Intelligibility estimation: A perceptual study in which listeners were asked to rate the intelligibility of sentences and conversational samples produced in Experiment 1.

Chapter 4 reports the results of these experiments, and Chapter 5 concludes by interpreting these results in the context of existing literature while addressing study limitations and future directions.

2 Background

2.1 Parkinson's disease

PD is a neurodegenerative movement disorder that affects approximately 1% to 3% of people over the age of 60, making it the second most common neurodegenerative disease following Alzheimer's (De Lau & Breteler, 2006; Nussbaum & Ellis, 2003; Wirdefeldt, Adami, Cole, Trichopoulos, & Mandel, 2011). The cardinal motor symptoms of the disease include bradykinesia (slowness), rigidity, postural instability, and resting tremor. Secondary symptoms may include speech disturbances, reduced facial expressions, dysphagia (swallowing disorder), micrographia (small handwriting), shuffling gait, and motor freezing (Jankovic, 2008; Wirdefeldt et al., 2011). Non-motor cognitive symptoms, such as anxiety, depression, and cognitive impairment, as well as dysautonomic symptoms may also occur (Wirdefeldt et al., 2011).

Though the etiology of PD is not fully understood, genetic and environmental risk factors have been identified (Nussbaum & Ellis, 2003; Wirdefeldt et al., 2011). Approximately 60% to 70% of PD diagnoses are considered idiopathic (Hughes, Daniel, Blankson, & Lees, 1992), and males are diagnosed approximately 1.5 times more frequently than females (Fahn, 2003; Wirdefeldt et al., 2011). Diagnosis of PD is made clinically, as there are no current definitive diagnostic tests (Postuma et al., 2015). Recently, there has been progress in the use of brain scan procedures such as DATScan (Thobois, Prange, Scheiber, & Broussolle, 2019). These show promise for identifying parkinsonism, but a major challenge at this time is distinguishing idiopathic PD from other differential parkinsonian disorders (i.e., multisystem atrophy, progressive supranuclear palsy; Thobois et al., 2019). Clinical diagnoses are typically based on the observed presence of the cardinal symptoms (specifically bradykinesia as well as rigidity and/or resting tremor) in the absence of other sources of neurological damage and accompanied by a positive response to dopaminergic replacement therapy (levodopa; Postuma et al., 2015). Post-mortem pathological criteria for a definitive PD diagnosis includes the presence of Lewy bodies in the brain (Gibb & Lees, 1988).

2.1.1 Neuropathology in PD

The primary parkinsonian symptoms develop as a result of dopaminergic cell loss accompanied by an accumulation of Lewy pathology and alphasynuclein protein in the brain (Nussbaum & Ellis, 2003; Poewe et al., 2017; Wirdefeldt et al., 2011). Dopaminergic depletion in the substantia nigra pars compacta leads to eventual loss in the striatum via the pathway of neuronal projections in the basal ganglia. This in turn leads to disruptions in the basal ganglia thalamocortical motor circuit, responsible for motor regulation, scaling, and initiation (Duffy, 2013; Poewe et al., 2017).

The first appearance of the cardinal motor symptoms (e.g., bradykinesia) typically only occurs after 50% of dopaminergic cells are depleted in the substantia nigra, and 80% in the striatum (Gonera, Hof, Berger, Weel, & Horstink, 1997; Wirdefeldt et al., 2011). More recently, earlier manifestations of the disease have been classified as prodromal and are thought to be related to nondopaminergic neuromodulators (Postuma, 2014; Postuma et al., 2015; Sapir, 2014).

A six-stage progression of PD proposed by Braak, Ghebremedhin, Rüb, Bratzke, & Del Tredici (2004) distinguished early "presymptomatic" and "symptomatic" phases. Each stage is marked by increased spread of Lewy bodies along a predictable neural topography from lower brainstem areas to higher cortical areas at more advanced stages. According to this model, the basal ganglia typically are not affected until Stages 3 and 4.

While Braak and colleagues termed the earlier stages (1 - 3) "presymptomatic," it is now generally acknowledged that these stages are consistent with the appearance of prodromal signs, including non-motor manifestations such as olfactory disturbances, sleep disturbances, and autonomic dysfunction (Berg et al., 2015; Gonera et al., 1997; Postuma, 2014).

Of particular relevance to this thesis is the finding that motor speech changes are also often present in these early stages, even though they may not be perceptually salient (Rusz, Cmejla, et al., 2013; Rusz, Čmejla, et al., 2013; Skodda et al., 2011; Tetrud, 1991). The pathophysiology of parkinsonian motor speech deficits is not fully understood at this time, but, given the increasing amount of evidence of voice and speech abnormalities that are detectable in prodromal PD, it has been proposed that these changes may occur in these "presymptomatic" stages (Sapir, 2014).

Further evidence for this hypothesis includes the observation that speech symptoms do not respond in the same way to dopaminergic replacement therapy as do the cardinal symptoms (Cushnie-Sparrow et al., 2018; Fabbri et al., 2017). One likely scenario is that speech production interacts in complex ways with both dopaminergic and nondopaminergic mechanisms, giving rise to the different responses observed in the speech and limb motor systems following treatment in PD (Cushnie-Sparrow et al., 2018; Im et al., 2018; Kompoliti, Wang, Goetz, Leurgans, & Raman, 2000; Skodda et al., 2013).

2.1.2 Treatment and management of PD

2.1.2.1 Pharmaceutical intervention

The most common intervention for the symptoms of PD is levodopa, a pharmaceutical treatment that may be administered orally or via a surgically implanted tube (i.e., Duodopa). Other medications used to treat PD in some instances include monoamine oxidase inhibitors, beta-blockers, or dopamine agonists (Connolly & Lang, 2014). Most individuals with PD demonstrate a robust response to levodopa, though prolonged exposure may eventually lead to a decline of the benefits of the medication (Obeso et al., 2010). This is often accompanied by the onset of motor fluctuations and involuntary movements, known as dyskinesias (Aquino & Fox, 2015). Some symptoms, such as speech and swallowing, show variable response to levodopa (Spencer, Morgan, & Blond, 2009). Early studies suggested such symptoms may also become progressively resistant to treatment over time (Bonnet, Loria, Saint-Hilaire, Lhermitte, & Agid, 1987; Klawans, 1986; Rascol et al., 2003). This resistance hypothesis has been recently challenged by findings suggesting that speech severity may mediate responsiveness of speech symptoms to levodopa (Cushnie-Sparrow et al., 2018; Im et al., 2018). According to these findings, more severe symptoms (e.g., poor voice quality, disfluencies) may demonstrate improvements under levodopa administration while more mild symptoms may deteriorate under the same conditions (Cushnie-Sparrow et al., 2018; Im et al., 2018).

2.1.2.2 Deep brain stimulation surgery

Individuals who experience a decline in the effectiveness or increase in adverse effects of levodopa may be candidates for an adjunctive surgical intervention known as deep brain stimulation (DBS; Limousin, Krack, & Pollak, 1998; Okun, 2012). Contraindications for DBS surgery include the presence of dementia or severe autonomic dysfunction, as well as atypical parkinsonism or symptoms that do not demonstrate a positive levodopa response (Okun, 2012). The surgery involves the implantation of permanent electrodes into a specific brain target. The most common neural targets are the subthalamic nucleus, globus pallidus, and, more recently, the pedunculopontine nucleus interna (Montgomery, 2007; Okun, 2012). The electrodes are connected to an impulse generator implanted in the chest under the skin, and they deliver constant electrical stimulation to the brain. The electrical current may be varied by alterations in the voltage, frequency, and pulse width of stimulation (Montgomery, 2007). Following DBS surgery, pharmaceutical treatment may still be prescribed, but the amount needed to manage symptoms typically is reduced or, in some cases, eliminated (Okun & Foote, 2004; Vingerhoets et al., 2002).

DBS is highly effective for treating many of the primary motor symptoms of PD (Deuschl et al., 2006; Krack et al., 2003; Limousin et al., 1998), though the specific mechanisms of action are not fully understood (Montgomery, 2007). The purpose of constant electrical stimulation is to modulate electrical activity in the brain thought to be responsible for the adverse motor symptoms of the disease (Okun, 2012). Control over the electrical field size and spread of stimulation is typically achieved through careful monitoring of the electrical parameter settings.

The effects of DBS on speech, unlike for the primary motor symptoms, are highly variable and often detrimental (Aldridge, Theodoros, Angwin, & Vogel, 2016; Iulianella, Adams, & Gow, 2008; Krack et al., 2003; Skodda, Grönheit, & Schlegel, 2012). The underlying causes of these detriments are not fully understood at this point, though the literature suggests relationships between speech and electrode localization (Montgomery, 2007; Tripoliti et al., 2014), pre-operative speech severity, longer disease duration (Tripoliti et al., 2014), and sub-optimal electrical parameter settings for speech (Abeyesekera et al., 2019; Aldridge et al., 2016; Chenausky, MacAuslan, & Goldhor, 2011; Farris & Giroux, 2013; Knowles et al., 2018; Skodda et al., 2012; Törnqvist, Schalén, & Rehncrona, 2005; Tripoliti, Zrinzo, & Martinez-Torresetal, 2011).

2.2 Hypokinetic dysarthria

This section will give a brief overview of the speech symptoms associated with PD and DBS. Sections 2.2.2 and 2.2.3 summarize PD-specific speech findings related to speech acoustics and intelligibility, while Sections 2.3 through 2.3.3 will more deeply address speech changes related to intelligibility and acoustics as observed as a function of changes to speech rate in PD as well as in healthy talkers.

Between 70% and 90% of people with PD will develop speech symptoms at some point during the disease (Logemann, Fisher, Boshes, & Blonsky, 1978; Mutch, Strudwick, Roy, & Downie, 1986; Müller et al., 2001). For most individuals with PD, these symptoms are consistent with a motor speech disorder known as hypokinetic dysarthria (HkD). *Dysarthria* refers to a collection of neurogenic motor speech disorders characterized by abnormalities in the strength, speed, range, steadiness, tone, force, or accuracy of speech movements (Darley et al., 1969b; Duffy, 2013). *Hypokinetic* reflects a down-scaling of oral speech movements. HkD has become essentially synonymous with the dysarthria of PD (Adams & Dykstra, 2009). In particular, seminal work conducted by Darley, Aronson, and Brown identified the most deviant perceptual features of the speech of people with PD to include abnormalities of articulation, namely imprecise consonants, abnormalities of rate, including short phrases, short rushes of speech, and variable rate, as well as other prosodic abnormalities including monopitch, monoloudness, and reduced stress (Darley et al., 1969b, 1969a). The authors labelled this cluster of symptoms prosodic insufficiency, which is thought to result from a limited range of movement, muscle rigidity, slowness, and reduced articulatory force (Darley et al., 1969a).

Logemann, Boshes, Fisher, & Siegfried (1973) sought to describe a typical profile of HkD based on physical rather than perceptual characterizations leading to the deficits described by Darley, Aronson, and Brown. Using the Fisher-Logemann Test of Articulation Competence (Fisher & Logemann, 1971) the authors assessed the co-occurrence of laryngeal and articulatory impairments involved in the speech of 200 people with PD. Voice abnormalities, which correspond in large part to laryngeal dysfunction, were found to be the most common symptom, present in 89% of patients. Approximately half of the patient group presented with phonatory deficits as the only speech symptom. Articulatory abnormalities constituted the second most common manifestation of vocal impairment, present in 45% of patients. All patients presenting with articulatory impairments also presented with co-occurring vocal impairments. The majority of articulatory abnormalities were related to posterior tongue involvement, with a subset of patients also demonstrating tongue blade impairments. Observance of this subset relationship led the authors to propose that vocal tract dysfunction in PD may progress in a posterior to anterior direction, affecting first the laryngeal function, followed by posterior and then anterior lingual control. This hypothesis, however, has been recently challenged in the face of findings that suggest posterior and anterior involvement even in mild disease stages (Read, Miller, & Kitsou, 2018). HkD, then, may be more accurately characterized by an overall down-scaling of articulatory movements that may affect all speech movements even early on in the disease (Read et al., 2018).

2.2.1 Speech changes associated with DBS

Further speech changes are often, but not always, seen following DBS surgery. As mentioned above, while DBS is highly effective at treating many of the primary motor symptoms of PD, speech outcomes do not demonstrate a similar benefit and, in many cases, speech symptoms worsen. Speech changes may be reported in reference to either an individual's pre-surgical speech, their speech when DBS is on versus off, or compared to a control group of people with PD who have not received the surgery. Reports of speech impairment following DBS surgery suggest tremendous variability in individual outcomes (Aldridge et al., 2016; Chenausky et al., 2011; Dromey & Bjarnason, 2011; Iulianella et al., 2008; Krack et al., 2003; Limousin et al., 1998; Skodda et al., 2012; Tripoliti et al., 2011). Primary findings will be summarized in this section.

Krack et al. (2003) showed that motor function impairments on the standard clinical PD rating scale improved for most outcome measures *except* for speech. A recent retrospective study also demonstrated that people who reported dissatisfaction with their DBS surgery outcomes tended to cite worsening of axial symptoms, including speech, as one of the principal reasons (Farris & Giroux, 2013).

Tripoliti et al. (2011) demonstrated that while people with PD with and without DBS showed an overall decline in intelligibility over the course of a year, the DBS group showed greater variability (including some individuals whose intelligibility improved), and a steeper overall slope of decline compared to the non-DBS cohort. Similar declines in speech intelligibility have been reported by other researchers as well (Plaha, Ben-Shlomo, Patel, & Gill, 2006; Rousseaux et al., 2004; Sidtis, Cameron, Bonura, & Sidtis, 2012; Törnqvist et al., 2005; Tripoliti et al., 2014; Tsuboi et al., 2014).

Greater impairments in articulation, as measured by acoustic metrics of vowel or consonant precision, have also been found following DBS (Dromey & Bjarnason, 2011; Eklund et al., 2014; Martel-Sauvageau et al., 2014, 2015; Putzer, Barry, & Moringlane, 2008; Sidtis, Alken, Tagliati, Alterman, & Van Lancker Sidtis, 2016). Other studies, however, have found relative improvements or no change in articulatory precision when DBS is on versus off (Dromey & Bjarnason, 2011; Karlsson et al., 2014, 2012; Tanaka et al., 2016) or after surgery (Åkesson, Lindh, & Hartelius, 2010). Other studies have reported improvements in voice quality symptoms following DBS as measured by voice acoustics (D'Alatri et al., 2008; Gentil, Pinto, Pollak, & Benabid, 2003) and perceptual voice ratings (Gentil, Chauvin, Pinto, Pollak, & Benabid, 2001; Zhou, Lee, Wang, & Jiang, 2009), though others have demonstrated declines (Klostermann et al., 2008; Tanaka et al., 2015; Tsuboi et al., 2014). The chosen speech task may play a role, with potentially greater impairment detected in spontaneous compared to repeated or read speech (Sidtis et al., 2012). This is consistent with reports of speech in PD in general demonstrating greater impairment often noted in more spontaneous speech production (Bunton & Keintz, 2008; Ho, Bradshaw, Iansek, & Alfredson, 1999; Sidtis, Rogers, Godier, Tagliati, & Sidtis, 2010).

Specific factors related to the DBS implant location and associated neural involvement are likely implicated in patients' responses as well. While there is a greater prevalence of studies that have explored the effects of speech following DBS of the subthalamic nucleus compared to other neural targets, those that have compared multiple targets have found differences in speech outcomes. DBS of the subthalamic nucleus may be associated with greater or more variable impairment compared to other neural targets such as the globus pallidus interna or pedunculopontine nucleus (Robertson et al., 2011; Tjaden, Greenlee, Brenk, Silverman, & Corcos, 2018), but not compared to the caudal zona incerta (Eklund et al., 2014; Karlsson et al., 2014). Left-unilateral subthalamic nucleus stimulation may be associated with worse speech outcomes as measured by mean fundamental frequency, articulatory accuracy, slower speech rate, and intelligibility compared to bilateral or right-unilateral (Santens, De Letter, Van Borsel, De Reuck, & Caemaert, 2003; Schulz et al., 2012; Wang et al., 2006).

Fenoy, McHenry, and Schiess (2016) explored the relationship between PD subtypes and the involvement of the dentatorubrothalamic tract, a fiber tract that has been suggested as being associated with speech deterioration. The authors demonstrated that greater dentatorubrothalamic tract involvement was associated with greater declines in spontaneous speech fluency, and tremor-dominant subtypes typically were less affected following DBS compared to akinetic-rigid types in terms of speech fluency and intelligibility.

Individuals with poorer preoperative speech intelligibility, longer disease duration, and more medially-located electrode contact may see greater detriments associated with DBS of the subthalamic nucleus (Fenoy et al., 2016; Tripoliti et al., 2014).

In summary, DBS is associated with high degrees of variability in speech outcomes, with many individuals demonstrating greater speech impairment following the surgery.

2.2.2 Acoustic characteristics of hypokinetic dysarthria

In addition to the general perceptual speech characteristics used to identify HkD, speech features can be described through objective acoustic measures.

In general, acoustic studies of speech in PD have demonstrated that HkD is associated with reductions in segment durations (Flint, Black, Campbell-Taylor, Gailey, & Levinton, 1992; McRae, Tjaden, & Schoonings, 2002), speech intensity (Fox & Ramig, 1997; Ho, Iansek, & Bradshaw, 2001; Holmes, Oates, Phyland, & Hughes, 2000), reduced variation of fundamental frequency, abnormal voice quality (Gamboa et al., 1997; Holmes et al., 2000; Kent, Vorperian, Kent, & Duffy, 2003; Rosen, Kent, Delaney, & Duffy, 2006), and reduced acoustic distinctiveness in both consonant (Lam & Tjaden, 2016; McRae et al., 2002; Tjaden & Wilding, 2004) and vowel production (Lam & Tjaden, 2016; McRae et al., 2002; Rusz, Cmejla, et al., 2013; Skodda et al., 2011; Tjaden et al., 2013a; Watson & Munson, 2008; Weismer, Jeng, Laures, Kent, & Kent, 2001). The following sections will explore in greater detail the literature regarding speech rate, timing, and segmental distinctiveness, which are of chief concern to the questions addressed in this thesis.

2.2.2.1 Rate and timing of speech

To reiterate, HkD is characterized by abnormal and often faster rates of speech (Darley et al., 1969b). Compared to other dysarthria subtypes, HkD is the only one in which faster speech is sometimes seen. Acoustically, speech rate can be described in terms of the overall rate of an utterance (e.g., in words or syllables per minute, with or without pauses), as well as of durations of individual speech segments and pauses.

Despite anecdotal reports of faster speech in individuals with PD, relatively few studies have found this to be the case (Flint et al., 1992; McRae et al., 2002). Many studies have failed to find any objective rate differences between people with PD and age-matched controls (Connor, Abbs, Cole, & Gracco, 1989; Kleinow, Smith, & Ramig, 2001; Ludlow, Connor, & Bassich, 1987; Skodda & Schlegel, 2008; Tjaden & Wilding, 2004; Walsh & Smith, 2012; Weismer et al., 2001), while others have demonstrated that people with PD present with *slower* connected speech rates (Hsu et al., 2017; Martínez-Sánchez et al., 2016) and alternate motion speech rates (Dworkin & Aronson, 1986; Ludlow et al., 1987; Wong, Murdoch, & Whelan, 2011). *Acceleration* of speech rate has also been reported in PD (for example, over the course of reading a passage), even in the absence of overall group differences in speech rate (Adams, 1994; Skodda & Schlegel, 2008) or syllable repetition (Ackermann, Hertrich, & Hehr, 1995; Hirose, Kiritani, & Sawashima, 1982; Netsell, Daniel, & Celesia, 1975; Skodda, 2011). In a review of speech symptoms reported in PD, Adams and Dykstra (2009) suggested a prevalence of abnormally fast rates of approximately 10%. As such, fast rates may not often be evident at the group level but may manifest in a subset of people with PD and HkD.

There is also evidence to suggest that the perception of fast rate in HkD may be due in part to increased coarticulation or “blurred” acoustic contrasts thought to arise from increased coarticulation (Kent & Rosenbek, 1982; Tjaden, 2000b; Weismer, 1984b). That is, speech segment durations themselves may not be reduced relative to healthy talkers, but acoustic events that typically occur in faster speech of healthy talkers, such as acoustic vowel reduction and formant transitions, occur during typical speech production in PD. In the speech of neurologically healthy individuals, reduced acoustic contrasts are common consequences of faster or more casual speech (Byrd, 1994; Lindblom, 1963, 1990; Picheny, Durlach, & Braida, 1986).

Voice onset time (VOT) is a temporal measure of stop consonant production that reflects the timing between the onset of a stop consonant release to the onset of voicing of the following vowel. VOT is considered to be reflective of laryngeal and supralaryngeal coordination (Weismer, 2006). It is the primary perceptual cue for stop consonant voicing, with voiceless stops typically characterized by longer VOT compared to voiced stops. VOT also systematically differs across distinct places of articulation, with more posterior placements associated with longer VOT (Abramson & Whalen, 2017; Cho, Whalen, & Docherty, 2019; Lisker & Abramson, 1964).

Reports of abnormalities in English VOT in PD are inconsistent (Bunton & Weismer, 2002; Cushnie-Sparrow, Adams, Knowles, Leszcz, & Jog, 2016; Fischer & Goberman, 2010; Flint et al., 1992; Forrest, Weismer, & Turner, 1989; Lieberman et al., 1992; Miller, Green, & Reeves, 1986; Weismer, 1984b). Forrest et al. (1989) found that voiced bilabial stops had longer average VOT for speakers with PD, making them more like voiceless stops, but did not find differences in voiceless bilabial VOT. On the other hand, other authors have found shorter *voiceless* VOT in talkers with PD (Flint et al., 1992; Weismer, 1984a). This has been attributed to stiffness in laryngeal musculature causing the vocal folds to have reduced abduction and preventing them to stay open as long as would be expected for typical voiceless VOT production (Weismer, 1984a). Others have demonstrated more *overlap* between voiced and voiceless VOT, calculated based on the distributions of both voiced and voiceless stops (Lieberman et al., 1992; Miller et al., 1986). Reports of no detectable differences in VOT between people with PD and healthy age-matched controls are also common in the literature (Bunton & Weismer, 2002; Cushnie-Sparrow et al., 2016), even when speech rate was controlled for (Fischer & Goberman, 2010; Ravizza, 2003).

2.2.2.2 Spectral acoustics

In addition to the temporal properties of speech, abnormalities reflecting the spectral quality of speech sounds are also present in PD. Much of this literature has focused on the spectral properties of vowel production. The most prominent acoustic-perceptual cue for vowel production is the frequency of two high-energy frequency bands known as the first and second formants (F1 and F2). Formant values can be measured for a given vowel, or as part of a composite measure to determine the spread of acoustic vowel space. Vowel space may be measured by the area formed by the polygon formed by a set of F1 and F2 values for multiple vowels, by the amount of dispersion that occurs from a vowel's F1 and F2 centroid, or by the acoustic distance between pairs of vowels.

Reductions in vowel space have been heavily documented in HkD relative to neurologically healthy speakers (Lam & Tjaden, 2016; Lansford & Liss, 2014; McRae et al., 2002; Rusz, Cmejla, et al., 2013; Skodda et al., 2012, 2011; Tjaden et al., 2013a; Watson & Munson, 2008; Whitfield & Goberman, 2014). A common finding across several of these studies, however, was non-statistically significant trends for reduced vowel space in HkD (Buccheri, 2013; Tjaden, 2003; Tjaden, Rivera, Wilding, & Turner, 2005; Weismer et al., 2001). This may be attributable, at least in part, to increased variability across speakers. More sensitive vowel space metrics designed to reduce interspeaker variability have since been applied to demonstrate differences in speakers with HkD (Fletcher et al., 2017a; Lansford & Liss, 2014; Sapir, Ramig, Spielman, & Fox, 2010, 2011; Skodda et al., 2012, 2011).

In addition to reduced overall vowel space, individuals with HkD have also demonstrated reduced vowel contrasts in front and back vowels, as indexed by the ratio of F2 in /i/ and /u/ (Rusz, Cmejla, et al., 2013; Sapir, Spielman, Ramig, Story, & Fox, 2007). Reduced second formant transitions, which reflect the speed and extent of tongue movement, have also been found (Feenaughty, Tjaden, & Sussman, 2014; Y. Kim et al., 2011; Kim, Weismer, Kent, & Duffy, 2009; Walsh & Smith, 2011; Yunusova, Westbury, & Weismer, 2005), though this finding is not uniform across all individuals or test words (Kim et al., 2009; Lam & Tjaden, 2016).

While considerably less attention has been given to spectral properties of consonant production in PD, consonant place distinction, as measured by the difference in spectral means, has been shown to be reduced in stops /t/ and /k/ (Lam & Tjaden, 2016; Tjaden & Wilding, 2004) and sibilants /s/ and /ʃ/ (McRae et al., 2002; Tjaden & Wilding, 2004). Tjaden and colleagues found that the reduction of the difference between spectral means for consonant pairs differing along a posterior-anterior place distinction was due to an overall lower spectral mean for the more anterior sounds (e.g., /s/ and /t/). The authors suggested that this might indicate a more posterior tongue position during speech production. Other studies, however, found no differences in spectral distinctiveness for sibilant (Lam & Tjaden, 2016; Yunusova et al., 2005) and stop consonant pairs (Cushnie-Sparrow et al., 2016).

Reports of spirantization are another feature of deviant acoustic stop consonant production in HkD. Spirantization refers to an abnormal amount of frication produced during stop occlusion, indicative of a “leaky” stop closure. Acoustically, this traditionally was measured as an increased visual presence of aperiodic spectral energy during the stop closure, but more objective acoustic metrics have since been applied. Spirantization has been reported in HkD (Canter, 1965; Kent & Rosenbek, 1982; Weismer, 1984b; c.f. Cushnie-Sparrow et al., 2016), as well as in greater amounts following DBS surgery (Chenausky et al., 2011; Karlsson et al., 2014).

Voicing through stop closure, particularly in voiceless stops when vocal fold vibration is expected to cease, is another speech feature reported to be more frequent in PD (Cushnie-Sparrow et al., 2016; Weismer, 1984b), and may reflect insufficient coordination between laryngeal and supralaryngeal gestures during stop production. It may also be the case, though, that this is a consequence of aging speech musculature rather than parkinsonism, as voicing through closure is also much higher in older compared to younger healthy adults (Weismer, 1984b).

Taken together, the literature that has explored timing and spectral features of HkD has demonstrated substantial variability, though evidence generally supports a hypothesis of articulatory undershoot.

2.2.3 Speech intelligibility in PD

Speech intelligibility refers to the degree to which a spoken utterance, that is, the *acoustic signal*, is understood by a typical listener (Miller, 2013; Weismer, 2008; Yorkston et al., 1996b). The speech changes in HkD are often associated with declines in intelligibility (Miller et al., 2007; Tjaden et al., 2014b; Weismer et al., 2001). This section briefly describes considerations when measuring speech intelligibility in dysarthria.

Speech intelligibility must be considered in the context of the speech task, listening environment, and listeners (Yorkston et al., 1996b). Intelligibility is a relative measure of spoken message transference, and thus is differentially impacted depending on variables in these other domains (Monsen, 1983). Other perceptual measures not chiefly concerned with understanding a spoken message, such as speech *severity*, *naturalness*, *bizarreness*, and *acceptability*, and *comprehensibility*, are distinct concepts, but may be related to intelligibility (Kent, Weismer, Kent, & Rosenbek, 1989; Yorkston et al., 1996b).

Speaker considerations: Intelligibility for people with PD has been shown to demonstrate a weak correlation with disease severity, and it is not necessarily related to age or disease duration (Miller et al., 2007). In healthy talkers, speech intelligibility has also been found to be higher in females compared to males (Bradlow, Nygaard, & Pisony, 1995; Bradlow, Torretta, & Pisoni, 1996).

Speech task: Higher intelligibility is often reported for individuals with PD in more structured speech tasks, such as reading and repetition, compared to spontaneous speech (Kempler & Van Lancker, 2002; Kent, 1996; Weismer, 1984b; Yorkston & Beukelman, 1981a), though this has not always been found to be the case (Bunton, 2008; Tjaden & Wilding, 2011a).

Listener considerations: While speech intelligibility refers to a listening task performed by a *typical* listener, what this means exactly may of course vary. For example, listeners who are considered truly *naive* and have had limited exposure to dysarthric speech may in theory behave differently compared to more experienced listeners, such as clinicians. The literature that has compared listener experiences, however, has generally found there to be no substantial differences related to expertise (Bunton, Kent, Duffy, Rosenbek, & Kent, 2007; Sussman & Tjaden, 2012). Similarly, other listener factors such as age, sex, and education have not demonstrated a differential effect on intelligibility scores (McHenry, 2011). Familiarization with the talker or speech samples, on the other hand, is associated with increased intelligibility in healthy speech (e.g., Bradlow et al., 1995) and dysarthria (DePaul & Kent, 2000; D'Innocenzo, Tjaden, & Greenman, 2006; Hustad & Cahill, 2003; Liss, Spitzer, Caviness, & Adler, 2002; Spitzer, Liss, Caviness, & Adler, 2000; Tjaden et al., 2005; c.f. Yorkston & Beukelman, 1983).

Listening task: The listening task and presentation are also important factors to consider when reporting intelligibility. For example, audio-only stimuli presentations have been shown to elicit lower intelligibility ratings compared to audio-visual presentations, where the listener has access to visual cues from the speaker (Hustad, 2006). Speech stimuli are sometimes mixed with multitalker noise to increase the difficulty of the task and avoid a ceiling effect (Bunton, 2006; Kuo, Tjaden, & Sussman, 2014; McAuliffe, Schaefer, O'Beirne, & LaPointe, 2009).

Intelligibility measurement considerations: Speech intelligibility may be measured through scalar estimation or identification (Hustad, 2006; Kent et al., 1989), sometimes referred to as *subjective* and *objective* measures, respectively (Yorkston & Beukelman, 1978). Estimation techniques may include percentage estimation (e.g., "What percentage did you understand?"), visual analog scale estimation, equally appearing interval scales, or direct magnitude estimation (Miller, 2013). Intelligibility identification measures usually refer to transcription or multiple choice, and this may be at the level of individual phonemes, whole words, or sentences (Miller, 2013).

Intelligibility estimation techniques are often considered more subjective compared to identification (e.g., transcription) because listeners' ratings depend on their own internal benchmark (Miller, 2013). Some scalar methods, such as fixed-modulus direct magnitude estimation, attempt to reduce this subjectivity by having listeners provide ratings with respect to a pre-determined standard, or modulus (Weismer & Laures, 2002). Despite this observation, scalar and transcription measures have demonstrated comparable inter-rater reliability (Enos, Abur, & Stepp, 2018; Stipancic, Tjaden, & Wilding, 2016; Tjaden & Wilding, 2011a) and rank ordering capabilities (Yorkston & Beukelman, 1978).

Another difference between estimation and identification metrics may be that scalar measures offer an estimation of the *magnitude* of the intelligibility deficit (Weismer, 2008), but may not be explanatory in nature (Bunton, 1999; Kent et al., 1989; Miller, 2013; Weismer & Martin, 1992). This is often treated as a criticism of estimation metrics. On the other hand, transcription metrics are thought to be more explanatory because they can provide information about the units of the speech signal that the listener did not understand. That is, transcription errors can provide insight into what components of the speech signal led to a breakdown in intelligibility (Weismer, 2008).

2.3 Speech rate modification

The goal of this section is to review the chief concerns behind the motivation for using speech rate modifications as a treatment strategy for individuals with HkD, and to outline the literature on the documented speech changes that occur in this group as well as in neurologically healthy talkers.

Improving speech intelligibility is one of the primary goals of behavioural speech intervention for individuals with dysarthria (Duffy, 2013). A common intervention technique for improving intelligibility in dysarthria is to help an individual learn to slow down their rate of speech (Duffy, 2013; Yorkston, Hakel, Beukelman, & Fager, 2007; Yorkston, Hammen, Beukelman, & Traynor, 1990). Though a less commonly reported therapeutic goal, a *faster* rate of speech may also be suggested for some individuals with dysarthria (Dagenais, Brown, & Moore, 2006). Speech rate is an appealing treatment variable because it is highly modifiable (Blanchet & Snyder, 2009; Yorkston, Dowden, & Beukelman, 1992), and rate reduction has successfully been demonstrated to improve speech intelligibility across multiple motor speech disorders (Yorkston et al., 2007), including HkD (Adams, 1994; Downie, Low, & Lindsay, 1981; Hammen, Yorkston, & Minifie, 1994; Hanson & Metter, 1983; LeDorze, Dionne, Ryalls, Julien, & Ouellet, 1992; Martens et al., 2015; Yorkston et al., 1990).

It should be noted that goal of rate reduction interventions is not necessarily a *normal rate* but rather *improved intelligibility*. Individuals with faster rates of speech may approach a more “normal” rate as a consequence, but the majority of speakers with dysarthria who demonstrate already slower than average speaking rates will use a rate of speech even less like that of healthy speakers. Nevertheless, there is likely a trade-off that will be exhibited, such that speech that is *too* slow may actually lead to worse intelligibility and/or reduced speech naturalness (Yorkston, Beukelman, Strand, & Bell, 1999). Considerably less literature has explored *faster* speech in talkers with PD (Kuo et al., 2014; McRae et al., 2002) or other types of dysarthria (D’Innocenzo et al., 2006; Turner, Tjaden, & Weismer, 1995), given that faster speech is rarely an appropriate therapeutic goal for these individuals.

In a review of the literature assessing rate, loudness, and prosody-based interventions for motor speech disorders, Yorkston et al. (2007) identified a need for better understanding of speaker candidacy for rate-reduction interventions, as well as better descriptions of how optimal rates are selected. A more complete understanding of the precise speech outcomes resulting from a wide range of rate adjustments across speech tasks and speaker profiles is needed in order to implement such findings in treatment.

HkD is one subtype of dysarthria that may stand to demonstrate considerable improvements in intelligibility following rate reduction methods (Yorkston et al., 1990). This may be the case regardless of whether an individual demonstrates a faster rate of speech, as slower speech is thought to be associated with larger oral movements, leading to greater acoustic contrasts and thus more understandable speech (Yorkston et al., 1999).

2.3.1 A note on the scope of this paper

Throughout this thesis, the concept of speech rate modification is considered in relation to a given speaker's typical rate of speech. That is, *slower* speech for a given speaker with a faster habitual rate of speech may actually approximate the typical speech of another person with a slower habitual rate, but it would still be considered *slower* for the first talker. It is assumed that habitual rates of speech will vary across individuals and speaking contexts, but the concept of speech rate modification here is considered with that in mind.

It should be noted that in this paper, speech intelligibility is considered as a reflection of the ability of a *typical listener* to understand a spoken utterance. Speech intelligibility is a relative measure of the transference of a spoken message depending on both the speaker and the listener, as well as the linguistic and environmental context in which the message is being conveyed (Monsen, 1983). Historically, a large body of literature has investigated aspects of speech production that enhance or decrease speech intelligibility for hearing-impaired listeners (e.g., Metz, Samar, Schiavetti, Sitler, & Whitehead, 1985; Monsen, 1983; Nickerson & Stevens, 1980; Stevens, Nickerson, & Rollins, 1983). While the contributions of this body of work are fundamental to our understanding of the relationship between acoustic aspects of speech production and understandability in this population and at large, for the purpose of this paper, speech intelligibility will be considered from the perspective of a typical listener with unimpaired hearing.

Another research direction that has gained much attention in recent years is the relationship between speech intelligibility and "clear speech" (e.g., Picheny, Durlach, & Braida, 1985; Krause & Braida, 2002; Tjaden et al., 2014). Clear speech is a style of speaking that talkers may use to communicate in difficult listening situations. It has also been investigated as a form of intervention for dysarthric speakers (Tjaden et al., 2013a, 2013b, 2014b; Whitfield & Goberman, 2014). While clear speech often is produced more slowly than conversational speech, other articulatory and prosodic factors are likely implicated in the intelligibility advantage associated with it. Evidence from the clear speech literature is reported here to supplement findings of speech rate-related changes where appropriate, though it should be noted that this mode of speaking is not the primary focus of the present study.

2.3.2 Speech rate and intelligibility

2.3.2.1 Possible explanations

Producing speech at a slower than habitual rate has been identified as an effective method for improving speech intelligibility in dysarthria (Adams, 1994; Downie et al., 1981; A. R. Fletcher et al., 2017b; Hammen et al., 1994; Hanson & Metter, 1983; LeDorze et al., 1992; Martens et al., 2015; Yorkston & Beukelman, 1981b; Yorkston et al., 1990). Several recent studies, however, have demonstrated that many talkers with dysarthria do not exhibit improved intelligibility when they reduce their speech rates, and some may even worsen (Fletcher et al., 2017b; Hall, 2013; Kuo et al., 2014; McAuliffe et al., 2017; Van Nuffelen, De Bodt, Vanderwegen, Van de Heyning, & Wuyts, 2010; Tjaden et al., 2004; Van Nuffelen, De Bodt, Wuyts, & Van de Heyning, 2009). In healthy talkers, it is not the case that people that have naturally slower habitual speech are necessarily more intelligible (Bradlow et al., 1996; Cox, Alexander, & Gilmore, 1987). Relatedly, while faster speech has received less focus in the literature, one study demonstrated that faster-than-normal speech is not necessarily associated with reduced intelligibility in talkers with PD (Kuo et al., 2014), and it may even be associated with increased naturalness or acceptability in some cases (Dagenais et al., 2006; Logan, Roberts, Pretto, & Morey, 2002; Sussman & Tjaden, 2012).

It is likely that the specific speech deficits of an individual impact their responsiveness to rate reduction (e.g., festinating speech versus hypophonia). A recent study demonstrated that baseline speech features could be used to predict whether loud or slow speech was a more appropriate treatment strategy for individuals with dysarthria (Fletcher et al., 2017b). The authors found that dysarthric individuals who had greater baseline speech imprecision and greater temporal variability in their vowel production were more likely to see greater gains in intelligibility in a slow speech condition. These findings supported previous research suggesting that more severe speakers may be more likely to benefit from slow speech (Hammen et al., 1994; Pilon, McIntosh, & Thaut, 1998).

The mechanism involved in *how* speech rate changes may bring about improvements or decrements in intelligibility is not yet fully understood. In general, the literature suggests that improvements in intelligibility related to slow speech cannot be easily explained by any one factor. Empirical evidence for the reasons underlying intelligibility improvements associated with rate reduction, however, is still quite limited. Studies that began testing hypotheses for these mechanisms, and their criticisms, are outlined below.

2.3.2.1.1 Processing time for listener

One hypothesis is that slowed speech allows listeners more time to decode a distorted signal (Hall, 2013; Hammen et al., 1994; Nishio, Tanaka, Sakabibara, & Abe, 2011). Naturally slowed speech in dysarthria is associated with longer speech durations, more pauses, and longer pauses (Hammen & Yorkston, 1996; Tjaden & Wilding, 2011b). One way to test whether slower speech rates, either by virtue of a slower signal alone or the presence of more pauses, are associated with an intelligibility benefit, is to synthetically manipulate dysarthric signals and measure the intelligibility changes. Studies that have synthetically manipulated speech by slowing it down or adding pauses, however, have generally demonstrated that these actions alone do not improve intelligibility to the same degree as naturally slowed speech for speakers with dysarthria (Hall, 2013; Hammen et al., 1994), nor for neurologically healthy talkers (Amano-Kusumoto & Hosom, 2011; Gordon-Salant, 1986; Krause & Braida, 2002). One study did demonstrate that inserting brief pauses between words produced by dysarthric talkers was associated with 5% increase in intelligibility, however (Gutek, Rochet, Robin, Yorkston, & Beukelman, 1996).

Informed by earlier investigations of deaf speech (Maassen, 1986; Massen & Povel, 1984; Osberger & Levitt, 1979; Uchanski, Choi, Braida, Reed, & Durlach, 1996), Hammen et al. (1994) synthetically altered the speech of talkers with HkD in three ways: adding pauses, increasing speech durations, and both adding pauses and increasing durations. Hammen et al. (1994) also had participants naturally slow their speech to 60% of habitual rates using pacing software (Beukelman, Yorkston, & Tice, 1997). Results demonstrated that, when the audio was played for listeners, only the *naturally* slowed dysarthric speech was associated with better intelligibility compared to the habitual condition. The findings of this study suggested that pauses or temporal increases alone could not explain benefits of slow speech in dysarthria. Later studies of synthetically altered dysarthric speech replicated these conclusions (Dagenais et al., 2006; Hall, 2013).

2.3.2.1.2 Articulatory/acoustic undershoot

Imprecise articulation in dysarthria is hypothesized to be related to *articulatory undershoot* (Ackermann & Ziegler, 1991; Logemann & Fisher, 1981; McAuliffe et al., 2006a, 2006b; Weismer, 1984b), that is, articulatory positions that do not achieve the target placement due to restricted movement. Slower speech may allow speakers with dysarthria more time to reach articulatory positions needed to produce more understandable speech. This has been found in some kinematic studies of dysarthric speech and rate (e.g., Adams, 1994; Caligiuri, 1989). The relationship between kinematic speech alterations and speech intelligibility is not straightforward, however. For example, Forrest et al. (1989) found relationships between speech intelligibility and labial amplitude and velocity in three speakers with HkD. Weismer, Yunusova, & Bunton (2012), however, found a relationship with tongue but not jaw or lip speed in people with HkD. More recently, Kearney et al. (2018) found that, despite improvements in articulatory working space following a novel treatment, only one of five individuals tested demonstrated improvements in intelligibility.

A related hypothesis is that improvements in *acoustic* space, which are related but not always directly attributable to increases in articulatory space, provide the primary underlying reasons for improvements in intelligibility in slower speech. Adams (1994) observed that greater acoustic distinctiveness was qualitatively visible in the slowed speech of a patient with HkD whose intelligibility showed a marked improvement. Greater *coarticulation* in slower speech (Hertrich & Ackermann, 1995a; Tjaden, 2000a), may also allow listeners to identify segment or word boundaries more accurately (Fogerty & Kewley-Port, 2009).

In general, findings suggest that slower rates of speech are associated with greater acoustic distinctiveness for talkers with HkD (Adams, 1994; McRae et al., 2002; Tjaden et al., 2005; Tjaden & Wilding, 2004). While studies of neurologically healthy talkers involve investigations of habitual rates of speech, rather than speech rate modifications, a similar relationship has been found such that naturally slower talkers tend to exhibit greater acoustic space (Bradlow et al., 1996; Tsao & Iqbal, 2006). Studies that have explored rate modification in healthy talkers have demonstrated evidence of increased distinctiveness in slow speech and decreased distinctiveness in fast speech (Adams, 1993; Miller et al., 1986; Tjaden, 2000a; Tjaden & Weismer, 1998; Tsao et al., 2006).

2.3.2.2 Other considerations

The relationship between speech intelligibility and rate in hypokinetic dysarthria is complex, and likely due to a number of factors including the number and location of pauses (Hammen & Yorkston, 1996; Hammen et al., 1994), speech naturalness (Yorkston et al., 1990), and a speaker's habitual rate and speech characteristics (Feenaughty et al., 2014). Other factors that may play a mediating role in these could be related to speaker-specific characteristics, the way in which speech rate modifications are elicited, or the degree to which speech rate is modified.

2.3.2.2.1 Speaker-specific/variability speech rate considerations

Individual talkers may respond differently to how and how much they modify their speaking rate. Taking a within-speaker approach in order to address the issue of interspeaker variability, Feenaughty et al. (2014) found that habitual speech rate was associated with a moderate effect on intelligibility in five of 12 speakers with PD. The direction of the effect, however, varied across speakers. By applying a within-speaker approach, the authors found that the speakers who had a faster habitual rate to begin with tended to be more intelligible when they slowed down, whereas those who had a slower habitual rate benefited from faster speech. It is important to note, however, that individuals with dysarthria who demonstrate slower speech in general may benefit from even further slowing.

Compared to other speaking conditions such as clear or loud speech, slow speech may not necessarily optimize intelligibility. For example, McAuliffe et al. (2017) found that loud, rather than slow speech led to overall greater improvement in intelligibility in six speakers with either HkD or ataxic dysarthria, despite the finding that four of the six were most intelligible in the slow condition. Furthermore, neither speaking rate or nor loudness alone correlated with speech intelligibility (though acoustic vowel measures did).

Fletcher et al. (2017b) explored how baseline speech characteristics could predict intelligibility gains in talkers with and without dysarthria across loud, slow, and habitual speaking conditions. The authors found that of their 43 dysarthric speakers (23 of whom had PD), approximately a third did not demonstrate improvements in intelligibility in either the loud or slow conditions. Those who did improve in the slow condition were characterized by having more severe speech imprecision and greater temporal vowel variability at baseline.

Relatedly, some researchers have suggested that, due to the variability of individual presentations in dysarthria, grouping individuals by their perceptual speech features or baseline intelligibility rather than dysarthria type may be more informative (Kim et al., 2011; McAuliffe et al., 2017). A speaker's baseline intelligibility may play a role in explaining gains or deterioration in intelligibility following rate modifications (Kuo et al., 2014; Pilon et al., 1998). Pilon et al. (1998) demonstrated that rate reduction led to improved intelligibility in two of three speakers with spastic-ataxic dysarthria; both speakers demonstrated lower baseline intelligibility compared to the third participant. As is seen to be the case with healthy speakers, whose baseline intelligibility leaves little room for improvement, the third and less severely impaired speaker's intelligibility did not improve with the slower rate.

Interspeaker variability may also be related to speaker-specific characteristics. In healthy talkers, men have been found to demonstrate somewhat faster speech than women (Byrd, 1994; Jacewicz, Fox, O'Neill, & Salmons, 2009) as well as more phonological reductions (Byrd, 1994). Normal healthy aging leads to changes in speaking rates, such that older speakers demonstrate slower, more variable habitual rates of speech compared to younger speakers (e.g., Jacewicz et al., 2009; Amerman & Parnell, 1992; Goozee, Stephenson, Murdoch, Darnell, & Lapointe, 2005; Mefferd & Corder, 2014; Wohlert & Smith, 1998). Slower speech in older talkers is largely achieved by producing longer segment durations (Benjamin, 1982; Fletcher, McAuliffe, Lansford, & Liss, 2015; Jacewicz et al., 2009; Quené, 2008; L. A. Ramig, 1983; Smith, Wasowicz, & Preston, 1988; van Brenk, Terband, van Lieshout, Lowit, & Maassen, 2009; Verhoeven, De Pauw, & Kloots, 2004; Yuan, Liberman,

& Cieri, 2006), though there is evidence that older speakers also produce shorter VOT (Benjamin, 1982). Despite consistent findings of overall slower habitual rates of speech in older speakers, it is unclear whether older speakers are able to increase their rate to the same extent as younger speakers when directed to do so (Goozee et al., 2005; van Brenk et al., 2009).

2.3.2.2.2 Rate control methods

Altering one's natural rate of speech is not an easy task, and there are multiple techniques that have been developed to elicit speech rate changes. Common methods include the following. Voluntary rate control involves an individual modifying their own rate speech to a self-selected target when told to do so. Magnitude production is one example of this, in which a speaker may be instructed to speak at a rate that feels two times faster or slower than their normal rate of speech. As previously mentioned, increased speech rate is a far less common treatment target compared to rate reduction, and as such comparisons of methods of eliciting *faster* speech for talkers with dysarthria have not been reported to the author's knowledge.

"Rigid" methods of rate control sometimes refer to techniques that require a person speak one word at a time and may be facilitated by different pacing mechanisms. Some pacing techniques in the literature include the use of metronomes or pacing boards. Specialized software (PACER; Beukelman et al., 1997) has also been developed for this purpose.

Feedback methods such as visual oscilloscopic or delayed auditory feedback (DAF) facilitate rate alterations as speakers adjust their own rate to a stimulus target. Visual feedback involves seeing an acoustic representation of one's speech and modifying the rate in order to achieve certain predetermined criteria, for example, "fill the screen" (Blanchet & Snyder, 2009). DAF requires that the user wear a small device consisting of a microphone and earpiece. As they speak, the acoustic signal is played back to them and is delayed by a small amount (usually between 25 ms and 200 ms). The goal is that the talker must slow their natural speech in order to "catch up" with the distorted signal (Blanchet & Snyder, 2009). DAF is thought to require relatively little training for a patient to learn (Yorkston, Beukelman, & Bell, 1988), whereas visual feedback requires the speaker attend more directly to the stimulus. Greater durations of DAF (i.e., within the range of 25 ms up to 200 ms) are typically associated with slower rates of speech, though gains in speech intelligibility are typically reported within 50ms to 150ms (Blanchet & Snyder, 2009).

Early case studies suggested that slowed speech aided by DAF were especially beneficial to some individuals with severe speech deficits characterized by short rushes and festinating speech (Adams, 1994; Downie et al., 1981; Hanson & Metter, 1983), but less effective for speakers with more mild impairments or impairments not specifically characterized by festination.

Comparisons of different rate control methods and target rates have yielded differences in the efficacy of eliciting slower speech (Odlozinski, 1998; Van Nuffelen et al., 2010) and improving intelligibility or naturalness (Van Nuffelen et al., 2010, 2009; Yorkston et al., 1990). Odlozinski (1998) compared four rate control procedures including magnitude production, DAF, visual pacing, and auditory pacing (i.e., metronome) at multiple rates with speakers with PD and healthy controls. Magnitude production was found to be less effective in rate reduction compared to DAF and visual or auditory pacing techniques for individuals with PD.

Van Nuffelen and colleagues compared seven rate control methods and their effect on visual analog scale intelligibility ratings in speakers with dysarthria, including PD (Van Nuffelen et al., 2010, 2009). The methods examined included speaking slower on demand, three pacing techniques (pacing board, alphabet board, hand tapping), and three DAF conditions (50ms, 100ms, 150ms) during a reading passage. Compared to habitual speech rate, the authors did *not* find an overall increase in intelligibility for any of the rate control methods. They did, however, find clinically significant intelligibility improvements in five of 19 participants, two of which had HkD (Van Nuffelen et al., 2009). An extension of this study (Van Nuffelen et al., 2010) included more individuals with dysarthria with the goal of identifying differences across rate methods. The authors found improvements in intelligibility in approximately half (13 of 27) of individuals following rate reduction. Furthermore, maximal intelligibility in these individuals did not necessarily correspond to the rate reduction method that led to the greatest decrease in speech rate. That is, the slowest rate was not necessarily the one associated with the highest intelligibility, though with the exception of the DAF conditions, the authors did not systematically explore a rate continuum across the other metrics. These findings indicate a complex relationship between rate reduction and intelligibility, and further emphasize a need to understand patterns within individuals that would make them viable candidates for rate reduction interventions.

Yorkston et al. (1990) distinguished four kinds of pacing techniques along *presentation* and *timing* dimensions. In terms of presentation, *additive* methods presented one word at a time whereas *cued* methods presented a whole sentence at a time, and instead underlined each word. In terms of timing, *metered* pacing involved prompting the production of each word at equal time intervals, whereas *rhythmic* pacing involved presenting each word at a rate intended to be more naturalistic. The authors hypothesized that metered rate techniques would disrupt speech naturalness and, consequently, intelligibility. Contrary to expectations, however, the authors found that the largest gains in intelligibility occurred in metered rate control conditions. Furthermore, the authors found intelligibility gains at the sentence, but not phonemic level, indicating that more connected speech was subject to greater improvements in intelligibility following rate reduction.

In summary, the effectiveness of rate modification may depend on a talker's habitual rate of speech and the manner of elicitation. These factors should be carefully considered in all studies involving rate manipulation.

2.3.2.2.3 Degree of adjustment

The majority of studies that have explored rate modification have done so by eliciting only one or two different rates (e.g., a "slow" and/or "fast" rate). The extent to which a speaker slows down (or speeds up) likely has a bearing on the extent to which intelligibility changes are noted.

Two studies have demonstrated that slowing rates to 60% of speaker's habitual rate (compared to 80% and the habitual rate) led to substantial improvements in sentence intelligibility in speakers with HkD (Hammen et al., 1994; Yorkston et al., 1990).

In their review of different rate control methods, Van Nuffelen et al. (2009) explored three different DAF rates conditions: 50, 100, and 150 ms. As previously mentioned, greater durations of DAF (within the range if 50 ms up to 150ms) are typically associated with slower rates of speech. While there was a trend for the slower rates to be associated with *lower* intelligibility, differences between these DAF rate conditions were not statistically significant. Their follow-up study further found that maximal rate reduction was not necessarily associated with maximal intelligibility gains (Van Nuffelen et al., 2010).

Yorkston et al. (1999) described the likelihood of a trade-off between speech accuracy and speech naturalness such that, for a given speaker with dysarthria, there may exist an intelligibility peak. Speaking *too* slowly in relation to this hypothetical peak would result in poorer understanding because of compromised speech naturalness, whereas speaking too quickly would lead to imprecise articulation. Yorkston et al. (1999) asserted that the goal of speech rate modification intervention is to identify a target rate that "will allow an optimal level of intelligibility without degrading naturalness unnecessarily" (pp. 416).

Few studies have examined more than two rate manipulation conditions in dysarthria (Tjaden, 2000a, 2003), and none of these multi-rate studies explored changes in intelligibility. Evidence from multi-rate studies in healthy talkers suggests that individuals do not modify their speaking rate in a linear fashion (Adams, 1993; Tsao et al., 2006). Rather, healthy speakers tend to make smaller adjustments on the faster end of the rate continuum, and larger adjustments on the slower end, resulting in a quadratic or more complex non-linear relationship between intended and actual speech rate (Adams, 1993). The associated changes in intelligibility along such a continuum is presently unknown.

2.3.2.3 Summary

The relationship between speech rate and intelligibility is likely nonlinear, though the precise relationship is difficult to speculate without the inclusion of a wider range of rate conditions. Van Nuffelen et al. (2010) found that while improvements in intelligibility were observed in approximately half of their subjects following rate reduction, the maximal rate reduction was not necessarily associated with maximal intelligibility. Yorkston et al. (1999) suggested a trade-off between speech intelligibility and rate, such that at both very slow and very fast rates intelligibility suffers, while intelligibility gains can be identified somewhere in between. This curvilinear relationship may differ across individuals.

Taken together, these findings suggest that the relationship between speech rate and intelligibility is complex and varies across individuals. A deeper exploration into the acoustic changes associated with speech rate modifications and their potential impact on intelligibility will be discussed in the following section.

2.3.3 Speech rate and acoustic distinctiveness

Literature on the relationship between speech rate and acoustic distinctiveness largely supports the finding that slower speech is associated with increases the acoustic distinctiveness of vowels, as evidenced by acoustic vowel expansion, in both healthy talkers (Fletcher et al., 2015; Fourakis, 1991; Tjaden & Wilding, 2004; Tsao & Iqbal, 2006; Turner et al., 1995; Weismer, Laures, Jeng, Kent, & Kent, 2000) and talkers with HkD (Buccheri, 2013; McRae et al., 2002; Tjaden et al., 2005; Tjaden & Wilding, 2004). Though less researched, some studies have also demonstrated increases in consonant spectral distinctiveness (McRae et al., 2002; Tjaden & Wilding, 2004). Most studies reporting acoustic consonantal changes resulting from speech rate adjustments, however, have focused instead on segment durations rather than spectral changes. This section will summarize this literature for both individuals with PD and neurologically healthy talkers.

2.3.3.1 Consonant distinctiveness

Spectral stop distinctiveness (specifically /t, k/) and sibilant distinctiveness (/s, ʃ/) have been shown to increase with slow speech in some but not all cases (McRae et al., 2002; Tjaden & Wilding, 2004). The paucity of literature on consonantal changes across distinct speaking modes and speaker groups makes identifying a pattern of change following rate reduction challenging. McRae et al. (2002) found that a slow speaking rate did not bring about significant changes in spectral distinctiveness, despite a trend of increased spectral distinctiveness in slow speech and reduced distinctiveness in fast speech. Tjaden & Wilding (2004) found that spectral stop distinctiveness did improve in slow speech condition in some speakers with HkD, but loud speech led to improvements in a greater number of speakers.

2.3.3.2 Vowel production

A central focus of several studies investigating the acoustic implications of speech rate is the consequences on vowel formant production. As stated above in Section 2.2.2.2, vowels are typically characterized acoustically by two high-energy frequency bands, F1 and F2. Formant values, measured in Hz, may be measured in isolation or together as an index. Formant trajectories, for example, involve identifying the degree of change from two time points over the course of vowel production. This is often done for diphthongs, which are produced as "moving vowels" (Kent, Weismer, Kent, Vorperian, & Duffy, 1999). Composite measures are often used as a proxy of acoustic working space during vowel production, for example, in measures of vowel space, distance, or dispersion.

2.3.3.2.1 Vowel space

In general, larger vowel space areas have been found to be associated with slowed speech in dysarthria, and smaller vowel spaces with faster speech (Buccheri, 2013; McRae et al., 2002; Tjaden et al., 2005; Tjaden & Wilding, 2004), though often these changes fail to reach statistical significance despite consistent trends. For example, Tjaden & Wilding (2004) found that in a group of speakers with dysarthria secondary to PD or to multiple sclerosis, while there was an overall expansion in vowel space in a slow speech condition, only the multiple sclerosis group showed significant differences. Similarly, changes in vowel space have been found to be significantly related to rate reductions in other dysarthric groups such as speakers with amyotrophic lateral sclerosis (Turner et al., 1995; Weismer et al., 2000) and cerebral palsy (Hustad & Lee, 2008).

Tjaden and colleagues have demonstrated that enhancements in acoustic vowel measures occur with slow speech, but that other speaking methods such as clear and loud speech may bring about greater changes in some cases. Tjaden & Wilding (2004) found that slow speech was associated with greater gains in vowel space compared to loud and clear speech, but the greatest gains in F2 slope, the degree of increase in vowel frequency over time, were found with loud speech. Loud speech was also associated with greater gains in acoustic stop consonant distinctiveness and intelligibility. Tjaden et al. (2013a) found that *clear* speech led to overall greater differences in tense and lax vowel space and vowel distinctiveness (measured by dispersion and vowel lambdas) when compared to habitual, loud, and slow speech conditions. Buccheri (2013) found that acoustic distances between front and back vowels /i/ and /a/, as well as measures of vowel dispersion, increased in both clear and slow speech.

As with speakers with HkD, in neurologically healthy individuals, slower rates of speech are associated with greater acoustic vowel space (Fletcher et al., 2015; Fourakis, 1991; Tjaden &

Wilding, 2004; Tsao & Iqbal, 2006; Turner et al., 1995; Weismer et al., 2000), whereas *faster* rates of speech are characterized by vowel formant undershoot (Lindblom, 1963; Tsao et al., 2006). Vowel undershoot occurs when speakers fail to reach the articulatory target to produce a particular vowel. The result is what is referred to as vowel *reduction* or *centralization,* in which the acoustic output corresponds to a more neutral tongue position, causing the resulting vowel productions to be less acoustically and possibly perceptually distinct from one another. Seminal work investigating these processes was carried out by Lindblom (1963). Vowel reduction is also known to occur in unstressed or phonologically short vowels, regardless of speaking rate (Moon & Lindblom, 1994), as well as in more casual, conversational speaking styles (Byrd, 1994; Lindblom, 1990; Picheny et al., 1986).

While this phenomenon of vowel formant undershoot in healthy talkers is a well-accepted occurrence in spoken language, both at habitual rates and when speaking rate is voluntarily adjusted (Bradlow et al., 1996; Fourakis, 1991; Miller, 1981; Turner et al., 1995), it is not ubiquitous (Engstrand, 1988; Gay, 1978; Hirata & Tsukada, 2004; Van Son & Pols, 1990). For example, Tsao et al. (2006) found that while slower habitual speakers tended to have larger vowel spaces than faster speakers, there was also substantially greater variability. As a consequence, significant differences in vowel space did not emerge between the “faster” and “slower” speakers. Neither Van Son & Pols (1990) nor Engstrand (1988) found evidence of vowel reduction in fast speech in Dutch or Swedish, respectively. Hirata & Tsukada (2004) found that, in Japanese, speaking rate did not have an effect on vowel formant space, though phonological vowel length did. Short mid vowels (/e/ and /o/) were most susceptible to reduction than longer more peripheral vowels.

2.3.3.2.2 Formant slopes, coarticulation, and rate

During speech, the tongue must move to achieve the distinct articulatory positions required to produce the intended sounds. The surrounding phonetic environment is known to influence the production of sounds as part of a process known as coarticulation (e.g., Fowler, 1980; Fowler & Saltzman, 1993; Rogers, 2014). In some cases, such as in fast or casual speech, these articulatory gestures overlap with one another, which has an effect on the acoustic output (Agwuele, Sussman, & Lindblom, 2008; Tjaden & Weismer, 1998).

Formant slopes capture the change in formant frequency over time and are thus inherently both 1) an ideal acoustic measure of coarticulation and 2) impacted by speech rate. Formant slopes are also thought to be an acoustic index that is directly associated with articulatory movement (Kent et al., 1999; Weismer & Martin, 1992; Weismer et al., 1992). In particular, the slope of the second formant (F2) relates to lingual advancement. Despite this, F2 slope measures have not received as much attention in speech rate modification studies compared to vowel space measures.

Coarticulation may manifest as perseverative (whereby the preceding sound influences the upcoming one) or anticipatory (anticipation of the upcoming sound influences the preceding one; Hertrich & Ackermann, 1995b). This is most often measured by the degree to which a given vowel formant transition is altered according the spectral characteristics of the surrounding phonetic environment.

In a series of graded speech rate experiments, in which participants were asked to produce a given phrase in a progressively faster or slower manner, Tjaden investigated coarticulation as measured by F2 trajectories in neurologically healthy talkers (Tjaden & Weismer, 1998) and talkers with PD and HkD (Tjaden, 2000a). She found similar coarticulatory patterns in faster and slower speech in talkers with HkD compared to healthy controls, but that these patterns were more systematic and less variable for the healthy talkers (Tjaden, 2000a, 2003).

Tjaden & Wilding (2004) measured F2 slope in diphthongs produced in loud and slow speech for talkers with dysarthria secondary to PD or multiple sclerosis. The authors found that F2 slopes were shallower (less precise) in slow speech compared to loud speech, and that a difference was noted for only half of the speakers. In healthy talkers, speech rate was found to partially account for the variability found in F2 trajectories in healthy talkers (taking into account not only F2 slope but also onset and target F2 frequencies), though there was high interspeaker variability (Tjaden & Weismer, 1998). In healthy talkers, *faster* speech is often associated with a greater degree of coarticulation compared to slow speech (Hertrich & Ackermann, 1995b; Tjaden & Weismer, 1998; cf. Van Son & Pols, 1990; Zsiga, 1994). Hertrich & Ackermann (1995b) found that, in German CVC contexts, slow speech was associated with reduced perseverative coarticulation, but unchanged or in some cases increased anticipatory coarticulation.

Taken together, the evidence for acoustic vowel articulation patterns occurring in fast and slow speech asserts the value of using appropriate vowel metrics to document speech changes that arise during rate modification. The evidence for consonant acoustics is less clear but suggests an important gap in the literature regarding how speech rate adjustments impact consonant precision and distinctiveness.

2.4 Relationships between speech acoustics and intelligibility

Research undertaken to identify the causal acoustic underpinnings of speech intelligibility in dysarthria has a rich history of methodological advancements and insights but faces many challenges in describing the overall relationship. It is currently unclear what the most reliable acoustic variables are of speech intelligibility deficits in PD, and how the relationship between speech acoustics and intelligibility changes across tasks, speech styles, and individuals. This section briefly reviews the literature that has investigated these patterns.

One difficulty is determining whether a relationship found between an acoustic variable and the perceptual measure of intelligibility is related to the "third variable effect" (Monsen, 1978; Weismer et al., 2001; Weismer & Martin, 1992), that is, whether the change in intelligibility relates to the variable in question or another related feature, such as severity. The use of multivariate regression may allow more meaningful interpretations of these relationships (Weismer & Martin, 1992). Measures of accuracy, rather than scalar metrics of intelligibility may also provide more insight into the causal relationships between articulatory-acoustic features and intelligibility deficits, which is of chief concern from a clinical standpoint.

Studies that have explored combinations of different speech subsystem variables in order to determine those that have the greatest influence on intelligibility point towards articulatory measures as potentially the greatest contributors, compared to, for example, voice measures (De Bodt, Huici, & Van De Heyning, 2002; Kim et al., 2011). Suprasegmental measures related to voice, resonance, and prosody have also been found to be related to intelligibility and acceptability in PD (Feenaughty et al., 2014; Kim et al., 2011; Martens et al., 2015; Whitehill, Ciocca, & Yiu, 2004).

In general, the literature supports findings that measures of vowel space and vowel overlap are often highly predictive of intelligibility in PD (Feenaughty et al., 2014; Kim & Choi, 2017; Kim et al., 2011; Lansford & Liss, 2014; McRae et al., 2002; Tjaden & Wilding, 2004). Measures of second formant (F2) movement, including F2 slopes in diphthongs (Kim et al., 2011, 2009; Tjaden et al., 2013b; Weismer et al., 2001, 2012) and F2 interquartile ranges across an utterance (Kuo et al., 2014; Yunusova et al., 2005) have also indicated strong relationships with intelligibility. Consonant metrics such as spectral differences between /s, ʃ/ and /t, k/ have been found to be related as well, but to a lesser extent than vowel articulatory measures (Kim et al., 2011; Tjaden & Wilding, 2004; Yunusova et al., 2005).

Kim et al. (2011) studied the contribution of eight acoustic measures on intelligibility in a group of people with dysarthria of varying etiologies including PD. That study found six measures to be significantly correlated across the pooled participant groups following a regression analysis. These were vowel space, F2 slope, sibilant spectral differences, voiceless interval duration, articulation rate, and F0 range. Speech intensity and a metric of variability were the two measures that did not demonstrate significant correlations with intelligibility. Within groups, only F2 slope was significantly associated with intelligibility for all dysarthric speaker groups. Articulation rate was significantly associated with all dysarthric groups *except* the PD group, which may reflect the speech rate abnormalities present in this population. This work suggests that acoustic correlates of speech intelligibility likely differ across distinct speech subgroups.

To this end, Yunusova et al. (2005) investigated acoustic correlates of intelligibility in PD at the group level, as well as within individual speakers. The authors found that across the group, F2 interquartile range and the number of words in a breath group were associated with listeners' abilities to identify spoken words and sentences. Within individual talkers, however, varied patterns of significant predictors of intelligibility emerged.

There is some evidence to suggest that the relationship between perceptual and acoustic variables varies at different speaking rates (Kuo et al., 2014; McRae et al., 2002; Tjaden & Wilding, 2004), but the relationship between intelligibility and acoustics across rates is largely inferred. It is not currently known what acoustic measures are most important when a wide range of speech rates are taken into account. What is clear is that the connection between acoustic speech features and intelligibility in PD is obviously multifactorial and depends on aspects of the speakers, listeners, and linguistic and situational factors.

2.5 Summary and rationale for proposed study

In summary, speech changes associated with modifications to speaking rate are diverse in people with PD. While many studies have demonstrated improvements in intelligibility following rate reduction, others have found declines or no change. In a small number of studies, improvements in intelligibility and voice quality have been found in faster speech. Yorkston et al. (1999) suggested that there is likely a trade-off between speech intelligibility and rate, such that for a given individual with PD, the optimal speech rate for intelligibility may fall somewhere between very slow and very fast. The tremendous amount of variability in the speech presentations of people with PD and especially in those who have received DBS make it difficult to discern what this “optimal” rate might be, given that nearly all studies that have investigated the effects of rate modification on speech acoustics and intelligibility to this point have explored only one or two rate adjustments (e.g., “speak two times slower”).

Studies of rate manipulation that encapsulate a wider continuum of rates have the power to inform researchers on 1) the extent to which modifying rate impacts the intelligibility of an utterance, 2) the kinds of changes that occur when speech rate is increased or decreased, and 3) how variations in these changes affect speaker groups and across individual speakers. HkD is a subtype of dysarthria associated primarily with PD that offers particularly valuable insight into these changes given the characteristic speech symptoms (abnormal rates, smaller movements, decreased segmental distinctiveness and prosodic insufficiency). Furthermore, individuals with PD who have received DBS often experience further speech detriment, but they have largely been understudied in rate modification literature.

To reiterate, slowing down one’s speech rate may allow time for talkers with HkD to reach the intended articulatory targets, rather than undershooting as is often evidenced in their habitual speech. Greater articulatory space thus leads to an expanded acoustic space, as demonstrated by increases in vowel space, formant transitions, and spectral moment differences. Given the variability and occasional improvements in speech outcomes associated with faster speech, this too merits further investigation for this population. Determining the precise changes that occur for these speakers across a range speech tasks and speech rates will permit researchers to better understand the specific differences that lead to optimal speech intelligibility and how these targets are achieved.

Examining the nature of changes 1) within each rate condition, 2) in relation to an individual speaker's habitual rate, and 3) over the course of the full continuum is important to understand group and individual differences. To date, no study has systematically examined a rate continuum in speakers with dysarthria in this way.

2.6 Primary research questions and hypotheses

The purpose of the current study is to identify changes in acoustic distinctiveness and speech intelligibility in PD across a broad continuum of speech rate alterations and speech tasks in order to advance understanding of the effects of rate manipulation, which is a common form of speech intervention.

Overall hypothesis: A broad continuum of speech rate changes will be associated with non-uniform changes in speech acoustics and speech intelligibility, and the nature of these speech changes will differ by speaker group and speech task. Specific research questions are stated below, followed by hypotheses generated from the current literature.

1. What differences in terms of the range of self-selected speech rates exist across speaker groups (younger and older controls, people with PD with and without DBS) when instructed to modify their rate from very slow to very fast?

- **Hypothesis**: The healthy control groups will demonstrate a wider range of volitional speech rates compared to the clinical groups.

2. What are the acoustic-phonetic changes that occur in PD and control groups along a speech rate continuum?

- **Hypothesis**: Slower speech will be increasingly associated with increases in speech segment durations and acoustic distinctiveness, whereas faster speech will be associated with the reverse trend. The magnitude of change will be greater for healthy talkers.

3. How does such speech intelligibility vary in PD and control groups along a speech rate continuum?

- **Hypothesis**: Overall, slower speech will be judged as more intelligible than fast speech, and control groups will be rated as more intelligible than PD groups.

4. What differences in speech intelligibility exist across speech tasks along a speech rate continuum?

- **Hypothesis**: More natural speech will be associated with stronger rate effects compared to less natural speech.

5. What is the relationship between speech acoustics and intelligibility across this rate continuum?

- **Hypothesis**: Acoustic predictors of intelligibility will include variables that demonstrate sensitivity to changes in speech rate.

3 Methods

3.1 Overview

Three experiments were completed in order to address the research questions. Experiment 1 was a speech production study. Speech stimuli collected from Experiment 1 were used for acoustic analyses, as well as stimuli presented in the subsequent perceptual experiments, Experiments 2 and 3. Experiment 2 was a perceptual transcription task featuring one speech task elicited in Experiment 1. Experiment 3 was a perceptual intelligibility estimation task featuring two speech tasks elicited in Experiment 1. The procedures involved in these experiments are detailed in Sections 3.2 - 3.3. Outcome variables of interest for each experiment are reported in Section 3.4, and the statistical analysis procedure details appear in Section 3.5.

The study was approved by the Health Sciences Research Ethics Board at Western University (Appendix A) and the Lawson Health Research Institute (Appendix B). All participants provided informed written consent. All three experiments were conducted under the same ethics approval.

3.2 Experiment 1: Speech production experiment

3.2.1 Participants

Four participant groups were included in the final study: two healthy cohorts and two PD cohorts. At the study onset, a fifth clinical group was included (individuals with ataxic dysarthria). Due to difficulties in recruitment, however, this group was dropped and is not reported here. The final groups consisted of 1) younger healthy control participants (YC), 2) older healthy control participants (OC), 3) people with Parkinson's disease and dysarthria who were receiving standard pharmaceutical interventions (PD), and 4) people with Parkinson's disease and dysarthria who had undergone deep brain stimulation surgery (DBS). Participant demographics for the PD, DBS, and OC groups are reported in Table 1, Table 2 and Table 3. YC group demographics are described in the text.

Exclusion criteria for all groups included 1) history of speech or language impairments (aside from HkD in the clinical groups)[1], 2) history of any neurological disorder (with the exception of PD for the clinical groups), 3) self-reported inadequate vision or uncorrected vision for reading print, and 4) currently undergoing speech therapy.

Seventeen participants in the YC group (9 male) were recruited as students or alumni from Western University. Inclusion criteria for the YC group included: 1) speak North American English as a first language 2) self-report normal speech and hearing, 3) be between the ages of 18 and 35.

Nineteen participants in the OC group were recruited from the community. Inclusion criteria for the OC group did not differ from the YC group with the exception of age and hearing status. The age restriction for the OC group was 50 to 90 years. Furthermore, self-reported age-related hearing concerns were permitted, as were the use of hearing aids. Two participants were excluded on the basis of being native British or Irish English speakers and having moved to Canada as adults (thus, maintaining a non-North American English accent), leaving 17 participants (11 male) in the final group.

Parkinson's participants (both the PD and the DBS groups) were recruited through the Movement Disorders Centre at University Hospital in London, Ontario. In addition to meeting the same inclusion criteria as the OC group, participants in the two clinical cohorts were deemed eligible if they had 1) received a diagnosis of Parkinson's disease at least one year prior, 2) were stabilized on anti-parkinsonian medication (or via surgical settings), and 3) had demonstrated evidence of at least mild HkD, as identified by a neurologist on the Unified Parkinson's Disease Rating Scale (UPDRS). Charts for the DBS participants were not available to the researcher in all cases, however, and so some DBS participants were recruited without prior knowledge of their speech symptoms.

---

[1] Exceptions to this included reports of childhood stuttering, which two PD participants reported had reemerged with PD (PD01, PD03). One DBS participant reported having had seen a speech-language pathologist as a child but could not recall the reason. Two participants reported having had a transient ischemic attacks that did not result in speech disturbances several years prior (PD01 and PD14). Two control participants reported no history of speech, language, or neurological disorders, but presented with mild articulatory abnormalities (OC06 and OC10). Given they were both aware of these and did not attribute them to any disease or incident, data from these participants was retained.

Twenty-two participants were recruited in the PD group (18 male). Of these, two reported they spoke Dutch as a native language and had English as young children after moving to Canada (PD310, PD320), but reported growing up speaking North American English and were thus included in the final study. One participant reported speaking Spanish at home as a child (PD01) but received all schooling in English in Ontario and reported English as his dominant language. One participant grew up speaking English in Trinidad and reported moving to Ontario in his twenties (PD16). One participant reported having received spinal cord stimulation approximately six months prior.

Thirteen participants were recruited for the DBS group (11 male). One DBS participant was excluded on the basis of a self-reported history of stroke that resulted in speech changes, and was thus excluded, leaving 12 in the final group.

In total, 69 participants were recruited for the study: 17 younger healthy controls (YC; 9 male), 17 older healthy controls (OC; 11 male), 22 people with PD receiving standard pharmaceutical treatment (PD; 18 male), and 12 people with PD who had received DBS surgical intervention (DBS; 10 male).

3.2.1.1 Additional intake and demographic information

All participants with the exception of those in the YC group underwent additional testing that was not exclusionary. This included a 40 dB SPL hearing screening at 0.5, 1, 2, and 4 kHz (unless the participant was wearing hearing aids), and completion of the Montreal Cognitive Assessment (MoCA). Two OC, four PD, and two DBS participants reported wearing dentures[2].

All participants supplied basic demographic information. For all participants, this included (a) age, (b) sex, and (c) any self-reported speech, language, hearing, or neurological concerns. The clinical groups also reported (e) the duration since their PD diagnosis, (f) whether they had previously received speech-language therapy, and when, and (g) a list of PD-specific medications. All measures were collected via self-report. Deviant perceptual

[2] Two DBS participants reported they did not wear dentures but were planning to be fit for them.

characteristics listed in the tables below were determined by the author by listening to the conversational speech samples elicited during the habitual speech task. Characteristics were noted according to the recommended dimensions used by the Mayo Clinic dysarthria studies (Duffy, 2013), which were adapted from the original features described by Darley, Aronson, and Brown (1969a; 1969b; 1975). Note that these features were not determined during a standardized passage reading, as is often recommended (Duffy, 2013); this was due to time restrictions during the study.

Table 1: Demographic data for Parkinson's disease group (no DBS).

Participant	Sex	Age	MoCA	Years post-diagnosis	PD medications	Deviant perceptual characteristics	Mean intelligibility
01	m	60	29	12	Levodopa	monopitch, monoloud, quiet voice, breathy voice	96.2
02	m	65	18	14	ApoLevocarb	monopitch, monoloud, quiet voice, imprecise consonants, distorted vowels, wet voice	80.2
03	m	65	23	12	Levodopa	repeated phonemes, imprecise consonants, monoloudness, short rushes of speech, short phrases, fast speech	50
04	m	66	28	35	Levodopa	hoarse voice, breathy voice, monoloudness, monopitch, short rushes of speech, imprecise consonants, fast rate	66.3
05	m	73	27	7	Levodopa	quiet voice, short phrases, short rushes of speech, fast rate	95.1
06	f	67	30	10	NA	fast rate, short rushes of speech, breathy voice	97.1
07	m	72	29	9	Levodopa, Amantadine	imprecise consonants, breathy voice, increased pitch	96.9

Participant	Sex	Age	MoCA	Years post-diagnosis	PD medications	Deviant perceptual characteristics	Mean intelligibility
08	m	85	24	4	Levodopa	harsh voice, hoarse voice, breathy voice, imprecise consonants	87.2
09	m	56	28	25	Levodopa, Amantadine	strained-strangled voice, high pitch, imprecise consonants, short rushes of speech, forced expiration, fast speech	81.5
10	m	71	25	5	Levodopa	imprecise consonants, distorted vowels, high pitch, hyponasality	86
11	m	68	25	8.5	Pramipexole, Levodopa	breathy voice, hoarse voice, whispered voice, high pitch	92.5
12	m	72	24	15	Levodopa, Pramipexole	hypernasality, monopitch	96.3
13	m	62	26	3	ApoLevocarb	hoarse voice, breathy voice, imprecise consonants, distorted vowels, wet voice, short rushes of speech	87.2
14	m	90	24	10	NA	hypernasality, high pitch, imprecise consonants, distorted vowels, monopitch	91.9
15	m	70	28	2	Levodopa	short rushes of speech, imprecise consonants, high pitch, breathy voice	90.7
16	m	73	23	10	Levodopa	hoarse voice, imprecise consonants, breathy voice, monopitch	92.8
17	f	71	26	5	Levodopa	hoarse voice	97.4

Participant	Sex	Age	MoCA	Years post-diagnosis	PD medications	Deviant perceptual characteristics	Mean intelligibility
18	m	64	28	6	Levodopa	imprecise consonants, distorted vowels, short rushes of speech, monopitch, monoloudness	68
19	f	68	28	18	Duodopa	breathy voice, hoarse voice, imprecise consonants	78.3
20	f	73	25	30	Levodopa, Mirapex, Amantadine, Apo-Gabapentine	imprecise consonants, short rushes of speech, fast rate	90.3
21	m	64	28	8	Mirapex	monopitch, monoloudness, imprecise consonants, high pitch, breathy voice, hoarse voice	89.4
22	m	71	25	10	Levodopa, Pramipexole	imprecise consonants	90.6

Note: PD = Parkinson's disease; MoCA = Montreal Cognitive Assessment (out of 30). Two PD participants (PD06, PD14) were unsure of their current medication list, which are listed here as NA. Deviant perceptual characteristics for the PD and DBS groups correspond to features noted during the habitual conversational speech samples. Mean intelligibility corresponds to the mean intelligibility ratings for each participant during sentence production in the habitual rate condition, as judged by the listener participants (NB: this was not a standardized intelligibility assessment).

Table 2: Demographic data for the Parkinson's disease with DBS group.

Participant	Sex	Age	MoCA	Years post-diagnosis	Years since DBS surgery	PD medications	Deviant perceptual characteristics	Mean intelligibility
01	m	60	24	12	2	Levodopa, Amantadine	hoarse voice, breathy voice, monoloudness, monopitch, imprecise consonants, prolonged intervals	91.8
02	f	71	16	25	9	Levodopa	imprecise consonants, short rushes of speech, fast rate, breathy voice, hoarse voice	75.3
03	m	63	24	18	9	Amantadine, Levodopa	imprecise consonants, short rushes of speech, increased rate overall, high pitch, breathy voice, hoarse voice	60.9
04	m	73	20	12	4	Levodopa	strained-strangled voice, imprecise consonants, distorted vowels, prolonged phonemes	33.8
05	m	56	27	16	6	Levodopa	harsh voice, imprecise consonants, monoloud	88.8
06	m	59	16	13	5	Levodopa, Amantadine, Sinemet	imprecise consonants, high pitch, breathy voice	80.4
07	f	69	25	16	3	Levodopa	strained-strangled voice, breathy voice, audible inspiration, loudness decay	89.9
08	m	66	28	14	6	Levodopa	pitch breaks, flutter, breathy voice, hoarse voice, imprecise consonants, inappropriate silences	90.6

Participant	Sex	Age	MoCA	Years post-diagnosis	Years since DBS surgery	PD medications	Deviant perceptual characteristics	Mean intelligibility
09	m	55	28	8	1	Levodopa	imprecise consonants, hoarse voice, short rushes of speech, loudness decay, fast rate	82.5
10	m	66	23	4	3	Levodopa	high pitch, hypernasality, imprecise consonants, short rushes of speech, fast rate	66.4
11	m	60	25	12	4	Levodopa, Ropinirole	harsh voice, breathy voice, imprecise consonants	74.4
12	m	66	28	14	7	Levodopa	imprecise consonants, hoarse voice, breathy voice, short rushes of speech, fast rate	NA

Note: PD = Parkinson's disease; DBS = Deep brain stimulation; MoCA = Montreal Cognitive Assessment (out of 30). Deviant perceptual characteristics for the PD and DBS groups correspond to features noted during the habitual conversational speech samples. Mean intelligibility corresponds to the mean intelligibility ratings for each participant during sentence production in the habitual rate condition , as judged by the listener participants (NB: this was not a standardized intelligibility assessment).

Table 3: Demographic data for the older control group.

Participant	Sex	Age	MoCA	Mean intelligibility
01	f	71	30	97.2
02	f	76	29	97.5
03	m	77	25	89.6
05	f	71	30	97.1
06	m	80	21	92.7
07	m	80	28	97.8
08	f	69	28	NA
09	f	79	29	96.7
10	m	76	28	84.7
12	m	66	29	98.5
13	m	67	29	93.2
14	f	66	28	97.5
15	m	72	29	97.9
16	f	74	25	97.2
17	m	64	28	96.8
18	m	56	29	96
19	m	64	30	96.5

Note: OC = Healthy older controls (group). MoCA = Montreal Cognitive Assessment (out of 30). Mean intelligibility corresponds to the mean intelligibility ratings for each participant during sentence production in the habitual rate condition , as judged by the listener participants (NB: this was not a standardized intelligibility assessment).

3.2.2 Speech production data collection

All data collection took place in the Speech Movement Disorders Laboratory at Elborn College at Western University (London, ON). The clinical cohorts elected the time of day that they wished to come in to complete the study. In most cases, this coincided with their optimal ‘on’ state relative to their PD medications and self-reported fatigue. In some cases, participants were required to take their medication during the protocol and were permitted a break to do so[3]. The total time for data collection (including consent and additional data collection factors) was approximately two to three hours over a single visit.

3.2.2.1 Audio recording procedure

Recordings were made in an audiometric booth (Industrial Acoustic Company) using a 2017 15-inch Dell laptop computer (Inspiron 15). Participants wore a headset microphone (AkG c420), positioned 6 cm from the mouth, and connected to the laptop via a preamplifier and digitizing unit (M-Audio MobilePre) attached via USB. The headset was positioned so as to allow hearing aids and glasses to remain in place. Audio recordings were made via Praat (Boersma & Weenink, 2011) (for the signal calibration and for the final speech task, described below), or via a customized MatLab script (*MATLAB version 9.4.0 (R2018a)*, 2018), both of which digitized the audio signals at 44.1 kHz and 16 bits. Each experimental trial was saved automatically via the MatLab script as a separate .wav file. Participants' speech was also recorded using Praat for the purposes of practice (i.e., these recordings were not included in the analysis), via a lapel microphone connected to a 2014 MacBook Air via a preamplifier (Focusrite).

The following speech tasks were included: 1) audio signal calibration, 2) nonsense words in a carrier phrase, 3) sentence reading, 4) picture description, 5) conversation, and 6) maximum rate sentence reading. These are described in greater detail below. Calibration and the maximum rate task were recorded first and last, respectively. The remaining four experimental speech tasks were designed to elicit a continuum from more- to less-controlled tasks and were randomized according to the procedure described below.

---

[3] In one case, a participant in the PD group experienced a wearing-off effect of the medication that posed a problem with data collection. Specifically, he began to experience dyskinesias that were later realized to interfere with the recording quality, as the microphone repeatedly contacted his cheek. Two DBS participants chose not to complete the tasks due to fatigue, and therefore have incomplete data sets relative to the other participants. Specifically, DBS02 did not complete the fastest condition, and DBS04 did not complete the slowest condition and the two fastest conditions.

3.2.3 Calibration

The calibration procedure was carried out as detailed in Dykstra, Adams, and Jog (2015). Audio recordings for the calibration signals were made using Praat software (Boersma & Weenink, 2011) and were digitized at 44.1 kHz and 16 bits. Participants were instructed to produce a sustained "ah" vowel while the investigator positioned a sound level meter (SLM) 15 cm from the participant's mouth. The participant was encouraged to attempt to produce an "ah" of 70 dB (SPL-A; slow setting), which was indicated by the needle of the SLM reaching 0. In many cases, participants were unable to attain a steady 70 dB signal; in these cases, the investigator would transcribe the dB level when the signal was steady and call out the level that the participant did achieve by saying "stop." This procedure was carried out at least three times and was repeated any time the participant removed the headset microphone (e.g., during a break).

These calibration signals were then used to identify each participant's *calibration factor*, which would be linearly applied to the intensity of their speech signal in all subsequent analyses. The calibration factor was determined in the following way. The point in the sustained vowel where the investigator said "stop" and called out the intensity of the signal was located in the recording. A point 500 ms preceding this location was marked, and the average intensity of this 500 ms segment was measured in Praat. The difference between this measured intensity value and the actual intensity value was calculated for each of the three trials. The average of these three values was used as the *calibration factor* for all experimental trials that corresponded to that calibration session.

3.2.4 Experimental speech tasks

The speech tasks are described below in Sections 3.2.4.1 - 3.2.4.4. Each task was elicited once per rate condition. Speech rate conditions and the task randomization procedure are described below the task descriptions in Sections3.2.5 and 3.2.6.

3.2.4.1 Nonsense word in carrier phrase

A list of 52 disyllabic nonsense words designed to elicit minimal phonological and acoustic consonant and vowel contrasts was constructed. These words were elicited in the phrase "____. I'll say ___ again." The development of the word list was influenced by the

University Western Ontario Distinctive Features Differences Test (DFD; Cheesman & Jamieson, 1996) as well as by the word list used in the landmark study of English vowels by Peterson and Barney (1952). The 52 items in the present study were designed to contrast consonants and vowels in words of the form /əCVd/, where C represented one of 21 phonemic consonant sounds in the English language (Cheesman & Jamieson, 1996) and V represented one of the 4 corner vowels of English. Because plosives and sibilants[4] were of particular interest in the present study, a greater representation was included in the word list. Stops (/p, t, k, b, d, g/), sibilants (/s, ʃ, z/), and the voiceless glottal fricative /h/ appeared with each of the four vowels (/i, u, æ, ɑ/). The remaining 12 consonants (/f, v, θ, ð, tʃ, dʒ, m, n, l, ɹ, j, w/) appeared only with /ɑ/. The nonsense words were administered in four separate lists per condition, with each list containing a random selection of 13 items.

Prior to the experiment, participants were instructed to read aloud the list of nonsense words they would encounter. The purpose of this was to (a) familiarize the speakers with the words so that the habitual speech condition would not be the very first time they encountered the novel words (as in Vogel et al., 2017) and (b) to ensure target-like pronunciation of the words. During the experiment, if a word was clearly mispronounced due to the novelty of the words (and, i.e., not due to speech disturbances induced by the task or disease), the researcher (TK) would remind the participant of the target-like pronunciation that had been indicated in the word list reading.

Mispronunciations on certain words were not uncommon due to ambiguity of English orthography. In order to minimize this, strict spelling conventions were used in the nonsense word creation, and participants' attention were directed to this pattern if difficulty in remembering how to pronounce certain items persisted. Frequent mispronunciations included: 1) producing a voiced interdental fricative as voiceless (in fact, this occurred so frequently in spite of corrections, that this item was eventually discarded from the analyses), 2) producing a low front /æ/ as a low back or low mid /ɑ/ (e.g., "abad", "ahad" were both instructed to be pronounced as /əbæd/ and /əhæd/, but often were pronounced as /əbɑd/ and /əhɑd/). For the most part, participants were able to quickly acclimate to the target

[4] Sibilants were of interest at the study onset but are not reported here and will be the focus of future analyses.

pronunciations and did not need prompting. In some cases, however, the researcher needed to consistently remind participants of certain items. In these cases, when possible, the researcher attempted to prompt the participant that a word rhymed with something else, rather than say the word itself. If the participant was observed to intentionally hyperarticulate the following pronunciation after a correction, the researcher had them repeat it once more, and this last utterance was the token that was taken.

3.2.4.2 Sentence reading

A unique randomized list of six sentences was created for each participant and trial. Each list included words ranging from five to ten words in length (one sentence at each length). Sentences were extracted from the Sentence Intelligibility Test (SIT) item bank (Yorkston et al., 1996a). The SIT contains 1,100 sentences that range from five to 15 words in length (100 sentences at each length). Sentences are grammatically correct but semantically anomalous. Sentences were split into two short lists (5, 7, 9 words and 6, 8, 10 words) during task administration. A probe sentence, "She saw Pattie buy two poppies," was also included.

3.2.4.3 Picture description

To elicit more spontaneous but still controlled speech, participants described three simple picture scenes for each rate condition. Each scene was selected from the Diapix picture corpus (Baker & Hazan, 2011) in order to elicit keywords containing stops and sibilants. For this task, the examiner presented the participant with a high-quality colour 8.5"" x 11" print of each picture (one at a time) and prompted them to describe what they saw in the picture. Note that the analysis of this task was excluded from the final analyses but is described here to provide an accurate account of the speech task protocol.

3.2.4.4 Conversation

Participants engaged in approximately two minutes of spontaneous speech in which they were encouraged to talk about specific topics. The examiner presented the participant with one of seven topics typed out on an 8.5" x 11" sheet of paper. Participants were encouraged to first read aloud the prompt, then respond using at least a few sentences. Topics included: hobbies, favourite vacations, favourite foods, family, where you grew up, favourite books/TV shows, and what you do or used to do for work.

3.2.5 Speech rate conditions

Participants performed each speech task once for each of the seven speech rate conditions: habitual rate, three slower rates, and three faster rates. These were presented in blocks, with habitual always elicited first. Half of the participants in each group then performed the slow conditions next, (in order from least slow to slowest), followed by the fast conditions (least fast to fastest). The other half of the participants performed the fast block first. This counter-balanced, blocked design was used in order to control for the presence of task order effects, while allowing participants to adjust their speech rate magnitude in a sequential order.

The modified speaking rates were elicited using a combination of magnitude production (Tjaden & Wilding, 2004) and graded rate adjustments (Tjaden, 2000a). Magnitude production techniques are considered to elicit more natural speaking rate continua (Adams, Weismer, & Kent, 1993; Turner et al., 1995) and have been used in several studies of dysarthric speakers (e.g., Turner et al., 1995; Clark, Adams, Dykstra, Moodie, & Jog, 2014; Hall, 2013; Kuo et al., 2014; McRae et al., 2002; Tjaden et al., 2013b; Tjaden & Wilding, 2004). The graded nature of the task (each condition progressively slower or faster than the last), elicited by way of the rate blocks (slow versus fast), was chosen to facilitate adaptation to the modified rates. The blocked nature of the rate conditions allowed for rate adjustments similar to those elicited during a graded speaking task (Tjaden, 2000a), such that speakers were asked to grade their rate up or down within a block. Instructions were given in the spirit of direct magnitude production, but only the order of the rate blocks (slow, fast), not the individual rate conditions (2x/3x/4x) were randomized.

Participants were given the following instructions for each block:

> *Habitual (1)*: “Please say the following at your normal speaking rate.”
>
> *Slower conditions (3)*: “Please say the following at a rate that feels like 2x/3x/4x *slower* than your normal speaking rate. Try to slow your speech down by stretching out your voice, rather than pausing in between words.”

Faster conditions (3): "Please say the following at a rate that feels like 2x/3x/4x *faster* than your normal speaking rate, while trying to be as accurate as possible."

In addition to verbal instructions, participants had constant access to a visual stimulus comprised of a curved, numbered line (designed to look like a speedometer) and a movable arrow pointing to the target rate (Appendix C).

Prior to beginning a new rate condition, participants were provided with an opportunity to practice the new rate. The investigator, who was always seated beside them in the booth, presented them with the verbal and visual instructions, and provided them with a practice sentence ("Buy Bobby a poppy"). They were encouraged to read the sentence aloud at least two or three times, but also as many times as they needed to feel comfortable and accurate at the new rate. While the exact rate was not important, they were encouraged to at least be sure they felt faster or slower than the previous rate (depending on the block). They were recorded saying aloud these practice trials. Once satisfied that they had achieved the target rate during the practice, the investigator selected the most representative trial and played it back to them. This trial was also measured online to ensure that it was indeed faster or slower (as appropriate) compared to the previous condition. This sentence was then treated as a reference for the given condition and played back to the participants every five to ten trials in order to help them maintain their target rate. They could also request that it be played if they reported themselves having difficulty maintaining the target rate during the tasks[5].

3.2.6 Speech task randomization

Within each of the seven speech rate conditions, the speech tasks were presented in a quasi-randomized order. While the nonsense words, sentence reading, and picture description were never presented in the same order, the monologue task was always presented last. This was done in order to ensure that participants were maximally adjusted to the given target speaking

---

[5] Two exceptions existed to this pattern in the early stages of the study for young healthy controls. The first participant (YC101) was not given a reference, and the second participant (YC102) only heard the reference sentence it when she or the examiner decided she was veering away from the target rate. A stricter protocol was established afterwards. Later investigation of the data revealed that both participants mentioned here were successful in modulating their rate, and so they were kept in the final analyses.

rate by the time they were asked to engage in spontaneous conversational speech, in order to minimize the cognitive load of this task.

All speech stimuli were presented on a 15-inch Dell laptop computer via a customized MatLab script adapted from the McGill ProsodyLab template (Wagner, 2018) by the author (TK). For each trial, text appeared on the screen. In the reading tasks (nonsense words, sentence reading), participants were encouraged to read the text silently before they began speaking aloud. In the spontaneous speech tasks (picture description, monologue), instructions appeared on the screen, at which point the investigator would present the appropriate prompt. In all cases, the text on the computer screen would turn red to indicate that they could begin speaking. The investigator controlled the timing of when to advance to the next stimulus. Breaks were offered as needed.

At the end of all three rate blocks, participants performed one last task, in which they were prompted to read aloud a sentence as fast as they possibly could. The sentence for this task was the prompt sampled in each rate condition ("She saw Pattie buy two poppies"). Participants were permitted to read the sentence as many times as they wanted in order to reach their maximum rate. Upon reaching their maximum rate, the investigator prompted them to go even faster three more times in order to ensure that their maximum rate was truly obtained.

A subset of the speakers also repeated portions of the 2x-faster and 2x-slower conditions at the end for reliability purposes, though this task was eventually discarded for time reasons.

An example of the task schedule within a rate condition could be as follows:

> Picture 1, Nonce List 1, Pattie prompt, Nonce List 2, Sentences List 1, Picture 2, Picture 3, Nonce List 3, Sentences List 2, Nonce List 4, Conversation.

3.2.7 Speech rate

While the speech rate conditions (i.e., "Speak two times faster than normal") were designed to elicit a continuum of speech rates, it was anticipated that not all individuals would do this to the same degree. Of most interest to this study was how true rates of speech, regardless of condition, impacted speech intelligibility and speech acoustics. For this reason, the speech rate condition was treated as a means to achieve a continuum of speech rates. Speech rate was thus considered in the following ways.

- **Rate condition**: The rate condition in which the utterance was elicited (habitual, 2x, 3x, 4x slower/faster).

- **Actual speech rate**: Words per minute (WPM), calculated for reading tasks only. This was calculated for the sentence reading and nonsense word carrier phrases by dividing the number of words by the total utterance duration. Actual speech rate in words per minute was calculated for each individual utterance. This metric included within-sentence pauses if they occurred.

- **Mean habitual speech rate**: Average actual speech rate (in WPM) for each speaker and task, calculated from utterances elicited in the habitual condition reading tasks.

- **Proportional speech rate**: The proportion of the actual speech rate to the mean habitual speech rate. For each speaker, each utterance's actual rate was divided by that speaker's mean habitual rate. For example, an individual with a mean habitual rate of 200 WPM may have produced a slower utterance at 150 WPM, and a faster utterance at 250 WPM. The proportional rate of the slower utterance would be equal to $150/200$, or 0.75. The proportional rate of the faster utterance would be equal to $250/200$, or 1.25. Utterances spoken in a typical habitual manner for a given speaker should thus approximate 1, slower utterances are less than 1, and faster utterances are greater than 1.

In the present study, actual speech rates and its derivatives were used to analyze the nonsense word and the sentence reading tasks, while speech rate condition was used to analyze the conversation task.

3.2.7.1 Categorical treatment of proportional speech rate

Of primary interest in this study were changes that occur along a continuum of speech rates. Given the variation in actual speech rates, the proportional speech rates were the focus of the analyses. This would allow an interpretation of a significant difference between a "slow" and "habitual" or "fast" comparison to truly reflect differences in individual talkers' rates.

Treating speech rate as a categorical variable was thus also desirable, as it would allow a comparison of changes between distinct rates. To achieve this, the proportional rate values were binned into distinct categories, similarly to how the rate conditions were designed. It is thus expected, but not necessary, for the proportional rates to approximately map on to the rate conditions.

Based on visual inspection of the actual ranges produced across speakers, it was determined that a seven-step continuum, as per the rate-conditions, was too fine-grained to capture meaningful distinctions. A three-step continuum would have allowed a comparison of habitual, slow, and fast, but this was determined to be too *coarse* a distinction, as changes within slower and faster rates were evident. A binning procedure of five separate rate levels was chosen as the most appropriate technique.

It was important to not simply divide an individual's total range into five equal parts, but to instead divide the slower speech and faster speech into equal parts separately. The reason for this was that visual inspection demonstrated a nonlinear trend, such that talkers made larger adjustments in their slow speech than they did in their fast speech. This is also a pattern reported in the speech rate literature (Adams, 1993; Tsao et al., 2006).

To this end, proportional rate was binned into five separate levels (H1 = *habitual*; S2 = *slower*; S3 = *slowest*; F2 = *faster*; F3 = *fastest*) in the following way. For each speaker, the habitual mean served as the starting point. The range of values of their proportional speech rate was then split into two categories: slower and faster than their habitual mean. Each half was first divided into five equal bins. These were then collapsed into two and a half sections in the following way: the outer-most bins were collapsed into one ("slowest" or "fastest"), the next two bins were collapsed into another ("slower" or "faster"), and the final bin (which was equal to half of each of the other two) remained ("habitual"). The slow and fast sections were then combined, resulting in five sections. This ensured that the middle bin, corresponding to habitual rate, was centered relative to the slower and faster speech. Bin sizes on either side of habitual rate were not and were not expected to be equal in size.

Depending on the nature of a given speaker's rate distribution, the sizes and number of observations in each category were not equal for all speakers. Examples to illustrate this are presented below.

Consider Example 1 in Figure 1. This speaker had a mean habitual rate of 187 WPM and produced a fairly wide range of rates from 26 - 403 WPM (proportional range of 0.14 - 2.16). Example 2, in Figure 2, presents a speaker with a slower average rate of 171 WPM and a more restricted range of 77 - 210 WPM (proportional range of 0.45 - 1.23).

The binning procedure described here was chosen to ensure that each individual's range was taken into account, given the variability across speakers. Thus, when proportional speech rate is used as a metric, "slower" and "faster" do indeed refer to an *individual's* speech being slower or faster compared to their own *standard.* "Slower" for one individual could correspond to 80% of their habitual rate, but for another could correspond to 70%, depending on the range they produced and their habitual rate of speech. While bin sizes are not necessarily equal across participants, they reflect the true nature of that individual's speech rate modifications.

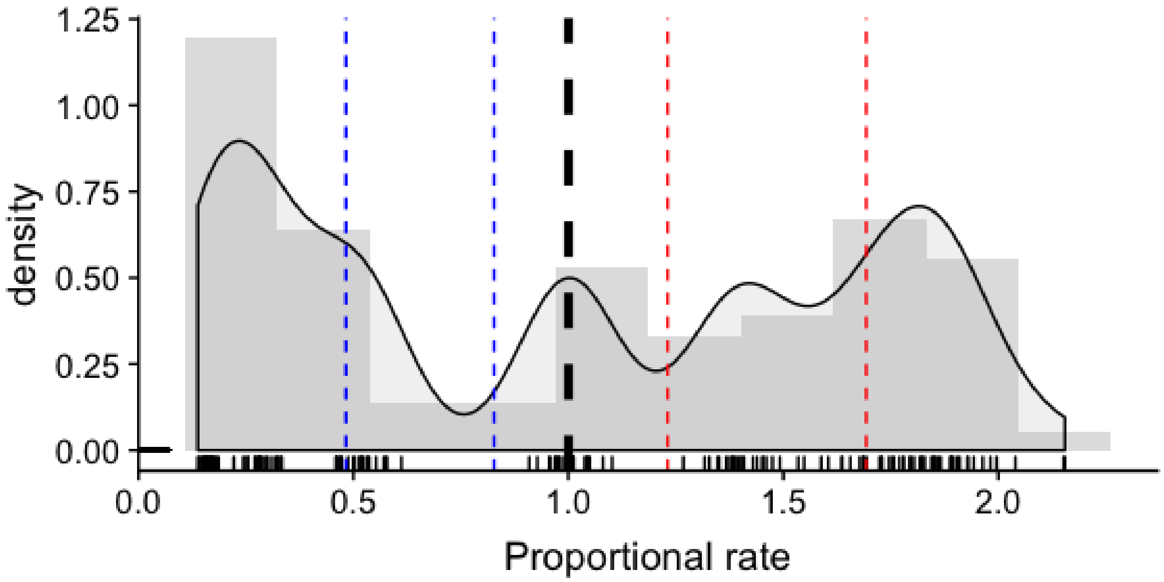


Figure 1: Example 1: Density plot smoothed with a Gaussian kernel showing the distribution of proportional speech rate production for participant PD301, overlaid with density histogram. The histogram bin width is 1/10 the range (0.216). Y-axis reflects the density of occurrence. This speaker with PD had a mean habitual speech rate of 187 WPM and produced a range from 26 WPM to 403 WPM. The black dotted line at x = 1 represents this speaker's habitual rate. Blue dotted lines represent their slower speech categorized into equally spaced slow rates (i.e., less than 1). Red dotted

lines represent their faster speech categorized into equally spaced fast rates (i.e., greater than 1).

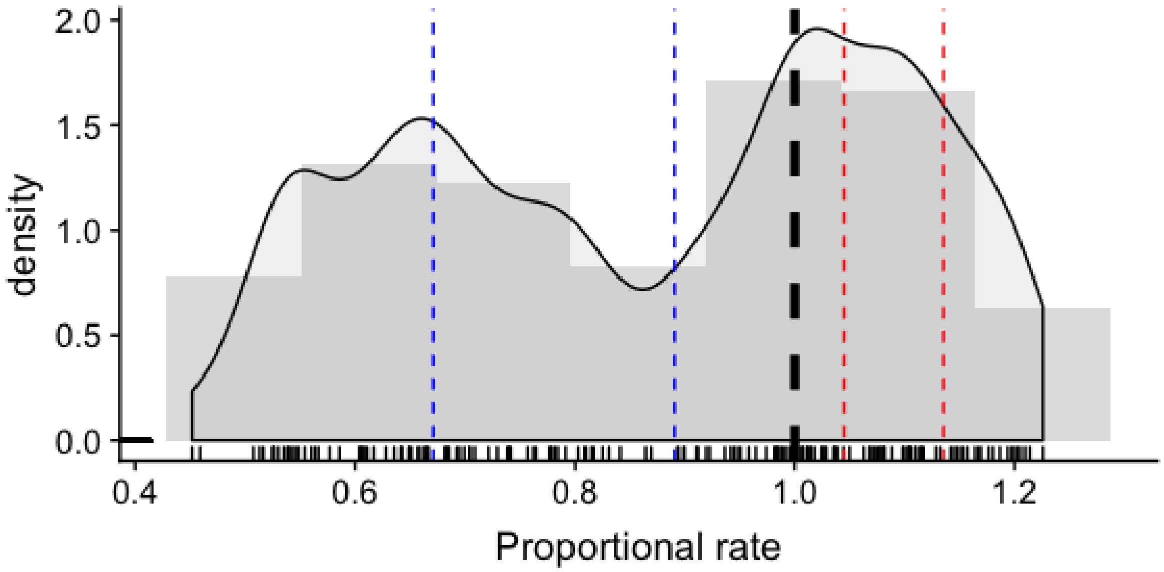


Figure 2: Example 2: Density plot smoothed with a Gaussian kernel showing the distribution of proportional speech rate production for participant PD312, overlaid with density histogram. The histogram bin width is 1/10 the range (0.123). Y-axis reflects the density of occurrence. This speaker with PD had a mean habitual speech rate of 171 WPM and produced a range from 77 WPM to 210 WPM. The black dotted line at x = 1 represents this speaker's habitual rate. Blue dotted lines represent their slower speech categorized into equally spaced slow rates (i.e., less than 1). Red dotted lines represent their faster speech categorized into equally spaced fast rates (i.e., greater than 1).

3.2.8 Acoustic analysis

The nonsense word utterances containing stop consonants were the focus of the primary acoustic analyses (items 1 - 24). This section describes the segmentation criteria and procedure. Specific acoustic outcome variables of interest are described in Section 3.2.8.1.5. Utterance extraction of the sentences and conversational speech, which were not subject to acoustic analyses, is described in Sections 3.3.2.2.

3.2.8.1 Manual segmentation

3.2.8.1.1 Automatic boundary labelling

Each utterance was first segmented at the utterance boundaries. Given that the carrier phrase was “Please say ___ again,” the utterance onset always corresponded with the /p/ release in “please,” and the utterance offset always corresponded with the offset of /n/ in “again.”

Once the utterances were extracted, they were automatically force-aligned using the Montreal Forced Aligner (MFA; M. McAuliffe et al., 2017). Utterances were aligned using the standard English acoustic models provided with MFA. Forced alignment was used as a tool merely to facilitate manual segmentation. That is, all segment boundaries were hand checked and adjusted as necessary. The original boundaries of this alignment procedure were saved for later comparison with manual alignment, though this will not be discussed here. Descriptions of the criteria for manual segmentation and adjustment are described in Section 3.2.8.1.3.

The standard output of automatic forced-alignment is a Praat TextGrid for each .wav audio file. The TextGrid contains word and phone boundaries that are time-aligned to the signal. Phones of interest in this study included the stop consonants and following vowels in the “aCVd” nonsense words.

3.2.8.1.2 Annotation protocol

A custom Praat script was written by the author to facilitate manual annotation of the speech segments of interest. The custom Praat script called “AutoVOT” (Keshet, Sonderegger, & Knowles, 2014), a software program for automatic detection of VOT.   The standard English classifier provided with the software was used to predict VOT. The output of AutoVOT is a TextGrid tier with the predicted VOT boundaries.

Each speaker was annotated one at a time, typically in a single session. Utterances for that speaker were presented in a fully randomized order as per the custom Praat script described above. A set of annotation codes were used to document the ambiguous cases described below, as well as any other observations. A typical speaker without any ambiguous cases comprised 168 audio files and took approximately 90 minutes to complete. Speakers with greater variability took longer, from between two to four hours to complete. The custom Praat script did the following:

- Took as input the force-aligned TextGrid and corresponding .wav file
- Called AutoVOT to predict VOT boundaries for the stop of interest, based on force-aligned stop boundaries +/- a 50 ms search range (to account for minor misalignment)
- Placed closure onset and vowel offset boundaries surrounding the AutoVOT predicted boundaries, based on the forced alignments
- Presented the modified TextGrid and .wav file for manual checking, correction, and note-taking in Praat
- Presented the vowel of interest (after the vowel boundaries had been corrected) with Praat's formant tracker turned on
- Allowed for adjustment of the formant values if necessary
- Saved all final results to an output file for later analysis

The output of the annotation procedure resulted in manual correction of the following acoustic events for the segments of interest:

- stop closure onset (which corresponded to the offset of the preceding schwa vowel)
- VOT onset, determined as the onset of the burst
- VOT offset, determined as the onset of periodicity in the following vowel
- vowel offset
- first and second formants of the vowel (measured at the 30 ms midpoint).

The preceding schwa vowel was not considered because of the frequency of schwa-deletion that made consistent boundary marking unreliable.

3.2.8.1.3 Manual boundary correction

A set of criteria were developed to maintain consistency in manual checking and correction of phone boundaries. All annotations were carried out by the author (TK).

Stop closure onset: Stop closure onset was determined by 1) a sharp decrease in amplitude in the preceding waveform, 2) a decrease in periodic complexity of the waveform, and 3) the absence of formant structure.

VOT onset: In most cases, the onset of VOT was identifiable by a clear burst. In many cases, however, the onset was more ambiguous, as is often reported in studies of VOT in clinical speech (Auzou et al., 2000; Fischer & Goberman, 2010; Karlsson et al., 2011). Informed by previous studies and patterns observed in these data, the following criteria were established.

- Multiple bursts were marked at the onset of the initial burst (Fischer & Goberman, 2010; Parveen & Goberman, 2014; Wang, Kent, Duffy, Thomas, & Weismer, 2004).

- Stops with clear frication preceding the burst were marked at the onset of frication present in the signal. This corresponded to "the transient with the strongest amplitude in the portion of the signal approximate to where an audible release was perceived" (Karlsson et al., 2014). These cases were also documented for later analysis which will not be discussed here.

- Stops with no obvious frication and no obvious burst could not reliably be marked as containing VOT.

A small subset of the stops could not be reliably marked as having a clear VOT onset. These were divided into three cases:

1. No VOT: 311 observations (3.7% of the data) had no obvious frication or burst; i.e., were unreleased (Özsancak, Auzou, Jan, & Hannequin, 2001). VOT was assigned a value of 0.001 seconds and closure duration was equivalent to the consonant duration. These special cases were omitted in the VOT analyses.

2. Complete frication: 172 cases (2.1%) had no clear distinction between the offset of the preceding vowel and the onset of frication; i.e., these stops were fully spirantized. In

these cases, VOT onset was considered as the onset of the consonant and offset of the preceding vowel. Closure duration was set to 0.001 seconds. These cases were omitted in the VDC analyses.

3. Completely omitted or glided: In a very small amount of cases (n = 37) there was no evident closure or release at all; i.e., the stop was unidentifiable. These cases were documented and removed from the analysis.

It should be noted that only positive VOTs were annotated. There were very few observed cases of prevoicing that would contribute to negative VOT. More frequently seen was either no voicing, partial voicing into closure, or complete voicing through closure.

VOT offset/vowel onset: VOT offset was determined as the onset of periodicity in the following vowel, marked on the part of the waveform crossing the x-axis going up. Two main causes of ambiguity were noted: quasi-periodicity in VOT, devoiced or breathy vowels, and voicing throughout closure. As such, the following criteria were followed:

- In the presence of quasi-periodicity, the onset of voicing was marked where there was an accompanying rise in amplitude in the signal. Praat’s pulse detection was also used to supplement particularly ambiguous decisions.

- In the presence of breathy or devoiced vowels, the offset of VOT was marked as an obvious visual change in the waveform and spectrogram indicating quasi-periodicity and formant-like spectral energy. Perceptual judgment was also used to supplement ambiguous cases.

Vowel offset: When possible, vowel offset was determined as the offset of periodicity and the onset of closure of the word-final /d/. In many cases, /d/ was unreleased or omitted, in which case vowel offset was marked using a combination of 1) visual inspection for changes in waveform and amplitude complexity, 2) changes in formant structure corresponding to a vocalic transition from the vowel of interest to the following schwa in “again”, and 3) audio perceptual judgments.

3.2.8.1.4 Vowel formant checking

The first two vowel formants (F1 and F2) were measured from a 30 ms section occurring at the midpoint of the vowel, using the boundaries established in the boundary correction phase described above. Formant values were manually checked using the same custom Praat script described above. Formant settings in Praat were uniformally set to begin, then set for each individual speaker on a case-by-case basis. Whenever possible, the same formant settings were kept consistent for a given speaker. Ambiguous cases were documented.

3.2.8.1.5 Final acoustic measures

The following acoustic measures were extracted or derived via other custom scripts, using the manually annotated segment boundaries as landmarks:

- **Voice onset time (VOT):** described above.

- **Voicing during closure (VDC)**: defined as the proportion of voicing during the stop consonant closure. A custom Praat script was written to calculate VDC using the manually annotated closure boundaries as time points, and the Voice Report feature in Praat (as described in Davidson, 2018). This feature extracts the "fraction of locally unvoiced frames", which was then subtracted in order to be converted to a proportion of voicing.

- **Quadrilateral vowel articulation index (QVAI)**: Vowel centralization was measured using a four-vowel articulation index (QVAI; Roy, Nissen, Dromey, & Sapir, 2009; Knowles et al., 2018; Sapir et al., 2011). This was calculated by averaging F1 and F2 for each of the four vowels for each speaker at each rate (using the rateProp bin categories). The following equation was used:

$$VAI = \frac{F2i + F2æ + F1æ + F1ɑ}{F1i + F1u + F2u + F2ɑ}$$

The QVAI measure has demonstrated greater sensitivity to acoustic vowel production in people with PD compared to traditional vowel space metrics (Sapir et al., 2011). Similar metrics involving a three-vowel index and its inverse, known as the formant centralization ratio, have also been used with dysarthric populations (Karlsson & van Doorn, 2012; Martel-

Sauvageau et al., 2014, 2015, 2015; Roy et al., 2009; Rusz, Cmejla, et al., 2013; Sapir et al., 2010; Skodda et al., 2011). VAI is designed to index vowel articulation in such a way that minimizes confounding effects of inter-speaker variability (Roy et al., 2009), thought to be one of the primary reasons for poor sensitivity of other metrics. In the formula described above, the numerator includes formant values that are expected to *decrease with centralization*, and the denominator includes formant values that are expected to *increase with centralization*. A larger QVAI thus reflects *less centralization* and greater expansion.

- **Vowel intensity**: Intensity was extracted for the entire vowel duration and was converted using the calibration factor procedure described in Section 3.2.3.

- **Harmonics-to-noise ratio (HNR)**: HNR was extracted from the vowel of interest across the entire vowel duration.

In summary, the five final acoustic measures of interest for Experiment 1 included: VOT, VDC, QVAI, Vowel intensity, Vowel HNR.

3.3 Experiments 2 and 3: Perceptual experiments

The audio recordings from Experiment 1 were prepared for two separate perceptual experiments (Experiments 2 and 3). Both will be described in more detail below in Sections 3.3.1 and 3.3.2, respectively. Briefly, Experiment 2 was a transcription task in which listeners were tasked with transcribing the nonsense words produced in Experiment 1, while Experiment 3 was a perceptual estimation task in which listeners rated how intelligible the more naturalistic speech samples (sentences and conversation) were using a visual analog scale (VAS). This section explains the methods for each experiment below.

3.3.1 Experiment 2: Transcription of nonsense words

3.3.1.1 Participants

Listeners were eight female second year speech-language pathology graduate students recruited from Western University graduate speech-language pathology second-year class. All were under the age of 35. All students received clinical motor speech hours for their time spent doing the tasks. Listeners passed a hearing screening at 20 dB SPL HL for octave frequencies from 250 to 8000 Hz[6]. All listening tasks were performed in a sound attenuated booth with audio stimuli presented via a pair of external speakers calibrated to 70 dB SPL.

3.3.1.2 Stimuli preparation

Speech stimuli from the nonsense word speech task described in Experiment 1 were presented to listeners in Experiment 2. To prepare the utterances for listeners, speech for each item/condition was extracted at the utterance boundaries +/- 50ms. These utterances were rescaled to 70 dB based on the intensity of the whole carrier phrase.

To increase the listening task difficulty and prevent ceiling effects in the read-speech conditions, the scaled utterances were then mixed with multi-talker babble. Mixing was performed using a standard multi-talker babble audio file (Audiotech – 4 talker noise) with a customized, modified Praat script (McCloy, 2013) at a signal-to-noise ratio of +3 dB. This is similar to noise levels reported in previous perceptual studies of dysarthria in order to reduce ceiling effects (Ferguson & Kewley-Port, 2002; Kuo et al., 2014; Maniwa, Jongman, & Wade, 2008; McAuliffe et al., 2009), though a SNR of +3 dB was specifically chosen based on pilot results preceding the experiment (others have reported, for example, -3 dB SNR, which was determined to be too low for these data).

3.3.1.3 Listening schedule

The transcription task was completed by each listener over the course of approximately four weeks. Listeners came in for approximately five two-hour sessions over this time-period (approximately 10 hours in total). The experiment at Session 1 was preceded by informed consent, demographic intake, a hearing screening, and a practice session and orientation to the task, all of which took approximately half an hour. All sessions therein were self-paced. Listeners were encouraged to take breaks as needed and only stay as long as they wanted for each session (as a result, some sessions were very short, i.e., completed over a lunch break, whereas most were approximately two hours long, and some were longer).

---

[6] One listener passed 250 Hz at 25 dB.

Each listener heard a random subset of five speakers from each of the control speaker groups (YC, OC), and a quasi-random subset of 10 speakers from the PD and DBS speaker groups (quasi-random in order to ensure that at least three DBS speakers were included for each listener despite the unbalanced group sizes), amounting in 20 speakers in total played to each listener.

Listeners heard all items[7] spoken by each listener in their playlist. Utterances produced by these twenty speakers were presented in a randomized order, with a randomized 20% of utterances repeated for reliability. This resulted in a playlist of approximately 8500 utterances for each listener[8]. The goal of this schedule was to ensure that the listeners underwent similar listening experiments with regards to variety of speakers and exposure to all elicited speech rates and items, while simultaneously minimizing the time requirements as much as possible.

A minimum of two listeners heard each talker. In some cases, because of the way the schedule was organized and the fact that not all talker groups were equal, three listeners were assigned to a given speaker. In these cases, the data from the third listener were discarded for the final analysis. This was done rather than provide shorter playlists for some of the listeners in order to ensure that the listener tasks were uniform.

3.3.1.4 Transcription instructions

Listeners were told that they would hear multiple speakers uttering the phrase "Please say ___ again," and that they were required to transcribe the word in the blank. They were informed that this word would always be a fake word of the form /əCVd/. To ensure consistency of orthographic representations of the nonce words, listeners were given an instructional sheet containing spelling conventions for the task. Consonants, they were told, could be any permissible consonant of English (with the exception of the voiced interdental

[7] 51 of the 52 utterances were included in the final presentation to the listeners. The voiced interdental fricative /ð/ item was dropped because most participants consistently had difficulty remembering how to pronounce it, despite prompting.

[8] In fact, a complete set would have included 8,568 utterances: 51 utterances x 7 rates x 20 speakers plus 20% repeated for reliability.

fricative /ð/[9]) and were provided with a list of these. Vowels were any permissible monophthong of English, also included on the list. The instructional list also contained several example words to help them with the task.

Listeners underwent a brief practice session under the supervision of the researcher to ensure that they understood the spelling conventions and the task. The practice trials included utterances spoken by the researcher in the form of the real test items. These practice items were also scaled to 70 dB. The second set of practice trials included the same items mixed in noise at +3 dB SNR (as were the real trials). Each item played once, and listeners were told that they could replay it up to one more time if they chose to (but were not required to do so).

3.3.1.5 Perceptual measures

Analyses for Experiment 2 focused exclusively on the transcription accuracy of consonants and vowels in the target words of nonsense words containing stops (items 1 - 24). These were the same items analyzed in Experiment 1. For each utterance and each listener, accuracy was logged for 1) the whole word (i.e., consonant and vowel), 2) on the consonant of interest, and 3) on the vowel of interest (Lansford & Liss, 2014). Consonant and vowel accuracy are of primary importance for this study. Each are treated as a binary response variable (correct or incorrect) and analyzed separately.

Each transcribed response was compared with the intended spoken target that had been elicited during that trial. Any response entered by listeners that did not correspond to the set of possible target responses given the orthographic criteria set for the listeners were manually checked and re-entered if necessary. Obvious errors were corrected (additional characters, omission of initial vowel or final consonant, etc.). Ambiguous responses were coded as *X* and counted as errors at either the consonant, vowel, or whole word level depending on the response (e.g., an answer transcribed "apiod" where the target response was "apid" would be scored as a /p/ for the *consonant*, but as an ambiguous error for the *vowel*, and thus would be scored as incorrect at the *whole word* level). Listeners were not permitted to leave any entries

---

[9] The voiced interdental fricative /ð/ was excluded due to concerns about consistent orthographic representation, and the fact that it had been intentionally excluded from the listening playlists.

blank, though they were instructed to type "NA" if they were completely unsure. In these cases, this was coded as an error at the word, consonant, and vowel levels.

To reiterate, in the final analyses, exactly two listeners heard each speaker. Listener responses were not averaged, but instead, *listener* was included in the statistical analysis as a parameter to account for expected variation[10]. This approach was chosen to 1) be able to model accuracy as a binomial (yes/no) variable; an average value would not allow this, and 2) retain as much of the data as possible, including variations across listeners. In other words, each speaker's token in the final analyses for transcription accuracy occurs twice; once for each listener that heard them, but the inclusion of *listener* as a covariate acts as a control for this in the final analysis.

Similar, though not identical, approaches to this statistical treatment of multiple listeners have been applied in other studies of dysarthria (Ferguson & Quené, 2014; McAuliffe et al., 2017). For example, Ferguson & Quené (2014) analyzed vowel transcription accuracy as a function of listener group (normal-hearing vs. hearing-impaired) for utterances produced by multiple talkers. In their case, they included *talker* (n=41) as a *random effect* in order to model the random effects of individual speakers. In the present study, including *listener* as a random effect was attempted first, but led to non-convergence in the model, presumably because of the small number of listeners (n=6 in total and only 2 per talker). Listeners were thus included as covariate fixed effects (i.e., independent variables) instead in order to account for inherent differences across them. This is a similar procedure to that described by McAuliffe et al. (2017), who included listener *group* (in their case, younger and older) as an independent variable (group was not relevant here because of the small number of listeners and the fact that listener behaviour was not a variable of interest). In the present study, listeners were expected to differ from one another, because each listener heard a difference subset of the data.

[10] Another viable approach would have been, rather than average, to err on the side of inaccuracy. That is, given two listeners per utterance, if either listener was incorrect in their transcription, the utterance would be coded as incorrect. This approach was not chosen in order to 1) avoid a floor effect) and 2) retain as much of the data as possible.

3.3.1.6 Inter- and intra-rater reliability

To calculate reliability, each response was compared with the intended spoken target that had been elicited during that trial. Both inter- and intra-rater reliability was calculated for each of the three accuracy categories (whole word, consonant, vowel) using Cohen's kappa (Cohen, 1960). Cohen's kappa was chosen because it is appropriate for binary categorical judgments and has been found to be robust to variations in listener experience and prevalence of stimuli (Grant, Button, & Snook, 2017). Because not all listeners heard the same subset of listeners, inter-rater reliability was calculated for every pair of listeners that *did* hear a subset of the same stimuli (n = 24 pairs), rather than across the whole set. A random sampling (n=100) of these data were used because not each pair heard the same number of overlapping utterances.

3.3.2 Experiment 3: Visual analog scale intelligibility estimation of sentences and conversational samples

3.3.2.1 Participants

Listeners were 6 female speech-language pathology graduate students who met the same criteria as in Experiment 2 (Section 3.3.1.1). Two of the listeners in Experiment 3 completed part of Experiment 2 during a pilot session. Given that these two tasks were quite distinct and did not contain the same stimuli, these listeners were kept for subsequent analysis.

3.3.2.2 Stimuli preparation

Sentences from the Sentence Intelligibility Task (3.2.4.2) were extracted at the utterance boundaries +/- 50ms and rescaled to 70 dB. The final data set contained 6 sentences per condition per speaker[11].

Spontaneous speech samples were extracted from the conversation task (described in Section 3.2.4.4) in the following way: when possible, between 10 - 20 seconds of continuous speech was extracted. Small pauses (e.g., less than one second) were considered acceptable. Some participants needed more prompting to remember to use the target rate during this task, and

[11] Sentences from 2 participants (OC208 and DBS512) were not included in the final playlist in error. Their data were included in the other listening tasks.

so 10-20 seconds of continuous speech was not always possible to extract. In these cases, one to three subsets of speech were identified and concatenated together until 10 - 20 seconds of speech were obtained. The final data set prepared for the listeners contained one sample per speaker. All audio files were rescaled to 70 dB.

3.3.2.3 Listening schedule

Listeners came in for approximately five sessions over a one-month period, as was the case for the listeners in Experiment 2. Listeners for Experiment 3 provided visual analog scale (VAS) ratings of speech intelligibility for the sentence production and monologue speech tasks. Unlike the transcription task, each listener in Experiment 3 heard all stimuli from all speakers. Ten percent of items were repeated for reliability purposes, amounting in approximately 3,600 utterances[12].

Utterances were presented in five blocks: four blocks for the sentence reading, and a single separate block for the conversational samples. Each sentence block contained all the utterances for four to five speakers from each group. The conversational block included all utterances from all speakers. Listeners typically completed one block per session. The blocks were presented in a different random order for each listener, and half of the listeners heard the conversational block first (the other half heard it last). Within each block, all utterances were completely randomized across all speakers and rate conditions (as with the transcription task).

The VAS tasks were administered via a customized Praat script written by the author that featured a horizontal line with anchors “Low intelligibility” and “High intelligibility” (demonstration in Appendix D). Listeners were instructed to rate the intelligibility of each utterance by clicking along the scale, which would place a thin vertical line. They were able to modify their rating until they were satisfied before moving on to the next trial. Listeners were instructed *not* to repeat trials (unless a there was a disruption in the room of some kind).

[12] A complete set would have included 3,665 utterances: six sentences plus one conversational sample each produced by 68 speakers at seven rates, plus 10% repeats.

3.3.2.4 Perceptual measures

Outcome measures for Experiment 3 included intelligibility estimates, represented as "% intelligible[13]", for the sentence reading and conversation tasks separately. Intelligibility ratings across all participants, tasks, and items were averaged across listeners.

3.3.2.5 Inter- and intra-rater reliability

Reliability of the speech intelligibility estimation task was calculated using the intraclass correlation coefficient (ICC; Koo & Li, 2016). Inter-rater reliability across the 6 listeners was examined using average consistency in a two-way random model (ICC 2, k) for each of the two tasks (sentences, conversation)[14]. Intra-rater reliability for each listener and task was examined using average agreement in a two-way mixed model (ICC 3, k)[15].

3.3.3 Relationship between speech acoustics and intelligibility

To address RQ5 (*what is the relationship between speech acoustics and intelligibility*), a final analysis of intelligibility was conducted in the following way. The outcome measure was transcription accuracy of the *whole word* from Experiment 2 (nonce words containing stop consonants). Because QVAI was one of the acoustic measures of interest, the data first had to be aggregated. The proportion of words correctly transcribed was calculated by aggregating the data over listeners, vowels, and place of articulation (PoA). QVAI was calculated on this aggregated data. The final data set for this analysis contained 1,949 observations. The final outcome measure was *proportion correct* measured at the word level.

---

[13] Percent intelligible is a slight misnomer, as the instructions to listeners were to rate intelligibility from *low* to *high* (in order to avoid ceiling effects of what it means to be "100% intelligible"), but is treated as a percentage value here for ease of interpretation of "percent along the visual analogue scale."

[14] The first two listeners were presented with more utterances for reliability purposes (20% instead of 10%). Given the length of the task, it was later decided that only 10% would be presented to listeners.

[15] Linear mixed effects regression was used in all cases except for non-convergence or singular fits. In such instances, an ANOVA was used instead. When calculating intra-rater reliability, non-convergence occurred for Listeners 4 (sentences and conversation) and 6 (sentences), and a singular fit was observed for Listener 1 (conversation).

3.4 Outcome measures

This section briefly summarizes the final outcome measures for each experiment.

3.4.1 Experiment 1: Acoustic variables

As outlined in Section 3.2.8, the following acoustic measures were of primary interest to the present study. Any transformations made for the statistical analyses are reported here.

1. **Voice onset time (VOT)**: VOT was treated as continuous and log-transformed to account for a right-tailed skew.

2. **Voicing during closure (VDC)**: VDC was treated as a binary categorical variable in the following way. Inspection of the data revealed that VDC ranged from 0 to 1, but approximately a third of the data in each group was equal to 1, indicating complete voicing through closure. As such, VDC was dichotomized into two categories: total voicing and some or no voicing through closure.

3. **Quadrilateral vowel articulation index (QVAI)**: QVAI was treated as continuous variable and was not transformed.

4. **Vowel intensity (dB)**: The calibrated intensity measure was treated as continuous variable and was not transformed.

5. **Harmonics-to-noise ratio (HNR)**: HNR was treated as continuous variable and was not transformed.

3.4.2 Experiment 2: Transcription

As described in Section 3.3.1.5, outcome measures for transcription included accuracy measures for:

1. **Consonant transcription accuracy**: (Stop) consonant accuracy was treated as a binary categorical variable (0 or 1).

2. **Vowel transcription accuracy**: Vowel accuracy was treated as a binary categorical variable (0 or 1).

Whole word accuracy was also measured and included as an outcome variable for the analysis of acoustic correlates of intelligibility.

3.4.3 Experiment 3: Visual analog scale estimation

As described in Section 3.3.2.4, outcome measures for the estimation task included **percent intelligible** for the sentence reading and conversational speech tasks.

Intelligibility was treated as a continuous variable. Early diagnostic plots of the models for both sentence and conversational intelligibility demonstrated that this outcome variable was highly left-skewed and violated assumptions of normality and homoskedasticity of residuals. The left skew indicated clustering of responses near 100% intelligibility, presumably due to a ceiling effect for some of the control participants. The intelligibility percent value was thus subtracted from a constant (100) and log-transformed.

3.4.4 Acoustic correlates of intelligibility

The outcome variable for the final analysis was **proportion of words correct** from Experiment 2 (i.e., taking into account both stop and vowel accuracy). *Proportion correct* was treated as a continuous variable ranging from 0 to 1. It was logit-transformed using the car() R package (Fox & Weisberg, 2011) with an adjustment factor of 0.2 in the final model in order to avoid proportions of 0 or 1.[16]

3.5 Statistical analysis

Differences in habitual rates of speech and speech rate ranges between each group, reported in Section 4.1.1 were calculated using Welch two-sample t-tests. Average differences between sentence and conversational intelligibility were compared using a Wilcoxon-signed rank test. All primary outcome variables described in Section 3.4 were modelled using linear or logistic mixed effects regression. The procedure for this mixed model analysis is described below for Experiments 1 - 3. The analysis to explore acoustic correlates of intelligibility

---

[16] This adjustment factor was chosen based on visual inspection of the distribution of the data as well as the residual plots in order to meet assumptions of residual normality and heterogeneity.

follows a similar procedure and is described at the end of this section. The final variables included in each analysis are detailed in the Results (Chapter 4).

3.5.1 Model building

All dependent variables were modelled as a function of the independent variables of interest (minimally speaker group and speech rate) using mixed effects regression with the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2018). Independent variables are herein referred to as *fixed effects*. A mixed modelling approach was chosen in order to allow for the inclusion of *random effects*, for example of participant or item. Random effects are able to at least partially account for the variability beyond that captured by the fixed effects. Mixed modelling[17] is a flexible, powerful tool for analysis that has recently been gaining popularity in the study of communication disorders (Harel & McAllister, 2019).

The p-values for the fixed effects terms in the linear mixed models were calculated using the Satterthwaite approximation from the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). For the logistic mixed models, p-values were calculated using asymptotic Wald tests (Bates et al., 2015). Post-hoc pairwise comparisons were computed using estimated marginal means (i.e., least-squares means) from the emmeans package, with p-values adjusted using the Tukey method (Lenth, 2018). One model for each dependent variable was constructed. Linear mixed effects modelling was used for continuous outcome variables, and logistic mixed effects were used for binary outcome variables. Logistic mixed effects modelling has been previously used to explore word and phoneme identification accuracy in dysarthria (Ferguson & Quené, 2014; McAuliffe et al., 2017).

All models (with one exception for RQ5) were constructed iteratively in the following way. The base model for a given outcome variable always included fixed effects of speaker group, speech rate (the binned proportional rates, e.g., H1, S2, etc.), and their interaction. The random effects structure included by-participant random intercepts and random slopes for

---

[17] Mixed models are sometimes also referred to as multilevel or hierarchical models (Harel & McAllister, 2019).

rate, as well as nested by-item random intercepts and random slopes for group. Random slope terms were uncorrelated in order to facilitate convergence and avoid over-parameterization of the models. This method of using uncorrelated random slopes has been reported in linguistics studies using mixed model procedures (Stuart-Smith, Sonderegger, Rathcke, & Macdonald, 2015; Tanner, Sonderegger, & Wagner, 2017).

To this base model, fixed effects terms of interest (depending on the outcome variable) were added one by one, starting with additional linguistic variables (e.g., voicing, place of articulation, vowel backness), then with speaker-specific variables where appropriate (e.g., sex). All two-way interactions were then tested, as well as all three-way interactions including group and rate. In order to limit the complexity and interpretation of the models, no four-way interactions or other three-way interactions were included.

At each stage, the new model was compared to the previous one using a likelihood ratio test implemented with the anova() function from the stats package (R Core Team, 2018). A term was kept if its inclusion led to an improvement in model fit, indicated by a smaller absolute Akaike Information Criterion (AIC) value and a p-value of less than 0.05.

The exception to this method was for the model pertaining to RQ5 (*what is the relationship between speech acoustics and intelligibility?*) which modelled intelligibility as a function of group, rate, sex, and all acoustic variables of interest in Experiment 1. This model included all variables of interest (i.e., did not undergo model selection) in order to evaluate the relative involvement of all acoustic variables were accounted for together.

3.5.2 Main effects and contrast coding

All categorical fixed effects terms were contrast coded in a manner that made theoretical sense for the levels being compared, as described below. These contrasts were consistent across models. Contrast coding is a form of centering categorical variables, which allows the intercept of the final model to be interpreted as the predicted value of the response when all predictor variables are held at their average values. It also allows for systematic interpretations of the predictor variables based on the specified contrasts themselves. For example, sum coding is a form of contrast coding that compares the two levels of a binary categorical variable against each other. Treatment coding specifies a particular reference level and compares all other levels to that reference. The number of contrasts for a given variable will always be the number of levels minus one. A binary variable such as voicing (voiced/voiceless) will thus have one contrast, whereas a five-level variable (such as the proportional rate variable) will have four contrasts.

All contrasting details are provided below for each individual variable.

3.5.2.1 Primary variables of interest: group and rate

Speaker group and speech rate were included in all models.

Speaker group was coded as a three- (in the acoustics models: OC, PD, DBS) or four- (in the intelligibility models, which additionally included the YC group) level variable and coded using reverse Helmert contrasts (as in Stuart-Smith et al., 2015). Helmert contrasts allow the mean of each level to be compared to the overall mean of the subsequent levels. The contrast scheme for group may be interpreted in the following way: 1) YC vs. Rest (i.e., OC, PD, and DBS groups combined); 2) OC vs. Clinical (PD and DBS combined); 3) PD vs. DBS.

Speech rate, using the proportional speech rate binning metric defined in Section 3.2.7.1, was coded using treatment contrasts with the habitual rate (H1) set as the reference level. That is, each contrast level compares one of the four modified rate bins to the habitual rate bin. Comparisons between each modified rate were discerned with post-hoc pairwise comparison testing.

3.5.2.2 Other variables of interest

The primary linguistic variables of interest in a given model varied depending on the outcome variable. This section describes the treatment of each variable entered in the models.

In Experiment 1, modelling different aspects of acoustic speech production, these variables were consonant voicing, consonant place of articulation, and vowel backness[18]. Experiment 2 models, which were constructed on the same data set as in Experiment 1, included the same

---

[18] Vowel backness, rather than vowel height was chosen as a potential predictor variable based on exploratory plots of the data demonstrating high degrees of back/front vowel distinctions that was not as present in high/low distinctions.

variables as above as well as *listener*. Experiment 3 modelled intelligibility estimation as a function of group, speech rate, and sentence length (in the case of the sentence reading task only).

Consonant voicing was sum-coded as a two-level categorical variable (voiced vs. voiceless). The *voiced* status was set as the reference level. Voicing was included in Experiments 1 and 2.

Consonant place of articulation was a categorical variable with three levels: bilabial, alveolar, and velar. PoA was coded using reverse Helmert contrasts (as was done for *Group*), such that the first contrast compared bilabial vs. non-bilabial (i.e., alveolar and velar), and the second contrast compared alveolar and velar. PoA was included in Experiments 1 and 2.

Vowel height was sum-coded as a two-level categorical variable (high vs. low) with *high* vowels (/i, u/) set as the reference level. Vowel height was included in Experiment 1.

Vowel backness was sum-coded as a two-level categorical variable (front vs. back) with *front* vowels (/i, æ/) set as the reference level. Vowel backness was included in Experiment 2.

Sentence length, i.e., number of words in a sentence (ranging from five to ten) was treated as an ordered six-level categorical variable ordered from shortest to longest[19] and coded using Helmert contrasts. This coding scheme allowed shorter sentence lengths to be progressively compared to longer sentences. The contrast scheme may be in the following way: 1) five vs. six-word sentences, 2) five and six vs. seven-word sentences, 3) five through seven vs. eight-word sentences, 4) five through eight vs. nine-word sentence, and 5) five through nine vs. ten-word sentences. Sentence length was included in Experiment 3.

One speaker-centric variable, ***speaker sex***, was also included. Sex was sum coded as a two-level categorical variable with *female* as the reference level. Speaker sex was included in Experiment 1.

---

[19] An initial attempt to code sentence length as a continuous variable was aborted because coding sentence length as a categorical variable in the manner described above led to lower AIC values, i.e., a better model fit. This is described in greater detail in the Results section.

As mentioned, ***listener*** was included and treatment coded with listener 1 as the reference level. This contrast, however, is not meaningful, and was only included to account for the variability across listeners in the model. Listener was included in Experiment 2.

One additional variable was included in the final analysis that modelled speech intelligibility as a function of the acoustic variables. This was ***audio clipping***, measured on the vowel, in order to account for any effects of a technical difficulty that was occasionally observed. Some speakers increased their volume over the course of some of the tasks, and in some cases, clipping was noted even though the microphone had been calibrated[20]. Clipping was measured on the vowel as a binary categorical variable (0 = no clipping, 1 = clipping) and was observed in less than 10% of stimuli overall. These utterances were included in the analyses rather than excluding them to avoid the intrusion of systematic bias. Clipping was used in the acoustic model of intelligibility.

Main effects terms and contrasts of interest for Experiments 1 - 3 are summarized in Table 4.

Table 4: Summary of main effects terms used in Experiments 1 - 3.

Variable	Contrasts	Experiments
Speaker group	1. YC vs. OC, PD, DBS 2. OC vs. PD, DBS 3. PD vs. DBS	1, 2, 3, ACI
Speech rate	1. S2 vs. H1 2. S3 vs. H1 3. F2 vs. H1 4. F3 vs. H1	1, 2, 3, ACI
Consonant voicing	1. voiced vs. voiceless	1, 2, ACI
Consonant PoA	1. bilabial vs. alveolar, velar 2. alveolar vs. velar	1, 2
Vowel height	1. high vs. low	1

[20] Unexpected behaviour in Praat also led to microphone signals being calibrated at a higher level than intended. In short, the recording software used for the calibration (Praat; Boersma & Weenink (2011)) was set to conduct stereo recordings even though the recording was carried out with a mono-channel USB preamplifier. The default setting in Praat for this procedure led to Praat not properly indicating when clipping occurred.

Variable	Contrasts	Experiments
Vowel backness	1. front vs. back	2
Speaker sex	1. female vs. male	1, ACI
Sentence length	1. 5 vs. 6-words 2. 5, 6 vs. 7-words 3. 5 - 7 vs. 8-words 4. 5 - 8 vs. 9-words 5. 5 - 9 vs. 10-words	3
Listener	(not meaningful)	2
Clipping	1. clip vs. no clip	ACI

Note: YC = younger healthy control group; OC = older healthy control group; PD = Parkinson's disease control group; DBS = deep brain stimulation group; H1 = habitual speech; S2 = slower speech; S3 = slowest speech; F2 = faster speech; F3 = fastest speech; PoA = place of articulation; ACI = acoustic correlates of intelligibility analysis. Note that only group contrasts 2 and 3 were used in Experiment 1, whereas all group contrasts were used in Experiments 2 and 3.

3.5.3 Random effects

Random intercepts take into account variability beyond that captured by the independent variables and allow variation by cluster (e.g., participants; Harel & McAllister, 2019). All models included by-speaker random intercepts. All but the QVAI model and the final acoustic correlates model included by-item random intercepts as well. Random slope terms account for variation of individuals within a cluster, for example, by accounting for different responses or slopes for individual participants. Where possible, all models included by-participant random slope terms for each contrast level of speech rate, and by-item random slope terms for each group contrast. There is some disagreement in the literature regarding how “maximal” a random effects structure should be (the reader is directed to Harel & McAllister, 2019 for a review). Over-specified random effects structures can lead to singular fits or non-convergence; the procedure detailed here was an attempt to mediate between accounting for important variation and not overfitting the models. Even so, in some cases the models did not converge, and the random slopes terms were reduced systematically. These cases, and any other deviation to this modelling procedure, are reported where relevant in the Results.

3.5.4 Acoustic correlates of intelligibility

The final model was not built using model selection, but rather included all the variables of interest. These were: *group*, *rate*, and the interaction between *group and rate*; all five *acoustic variables* from Experiment 1, as well as *consonant voicing*, and the interactions between *VOT and voicing* and *VDC and voicing*; *speaker sex*, and *audio clipping*. Vowel height/backness and place of articulation were not included because the data were aggregated over these variables to derive the QVAI variable. Random by-participant intercepts and slopes for rate were included.

4 Results

The organization of this chapter and the relationship of the analyses to the central research questions is as follows. RQ1 (*to what extent did the groups modify their speech rate?*) is answered first in Section 4.1.1. The acoustics results of Experiment 1 are presented in Section 4.1 in order to answer RQ2 (*what are the acoustic changes across groups and speech rates?*) along multiple dimensions, including consonant, vowel, and voice measures. The intelligibility results of Experiment 2 and 3 (Sections 4.2 and 4.3) follow in order to answer RQ3 (*what are the changes in intelligibility that arise across groups and speech rates?*) and a comparison of these experiments addresses RQ4 (*what are the changes in intelligibility that arise across speech rates and speech tasks?*). Finally, Section 4.4 revisits Experiments 1 and 2 in order to answer RQ5 (*what is the relationship between speech acoustics and intelligibility?*). Of primary interest for all of the research questions are the results of the group x rate interactions for each outcome measures.

Taken together, these results report on a complex relationship between speech rate, disordered speech, and other linguistic and contextual factors.

4.1 Experiment 1: Speech acoustics

4.1.1 Speech rate differences

4.1.1.1 Habitual speech rate

Before addressing RQ1 (*to what extent did the groups modify their speech rate?*), differences in mean *habitual* rates across the groups were first determined. Habitual speech rate was calculated as the mean rate in words per minute for each speaker in their habitual rate condition for the subset of items containing stop consonants (24 utterances per person). Mean habitual rates are reported in Table 5. The distribution of speech rates for each group is reported in Figure 3.

A series of Welch Two Sample t-tests were run to test for differences in habitual rates between groups. The YC group was found to have a significantly faster rate of speech compared to OC (p = 0.011) but did not differ from the PD (p = 0.831) or DBS (p = 0.863) groups. The OC group demonstrated significantly slower speech compared to the PD group

(p = 0.029) but did not differ from the DBS group (p = 0.11). The PD and DBS groups did not differ from each other (p= 0.983).

Taken together, this indicates that PD and DBS groups demonstrated a habitual rate of speech closer to that of younger adults than to their age-matched older healthy peers. The DBS group, despite having a near equal average rate to the PD group, did not significantly differ from the OC group, presumably because of greater variance.

Table 5: Mean habitual rates for each speaker group.

Group	N	WPM	StDev	StError	CI
YC	17	168.706	21.251	5.154	10.926
OC	17	146.809	25.663	6.224	13.195
PD	22	166.964	29.482	6.286	13.071
DBS	12	166.710	35.065	10.122	22.279

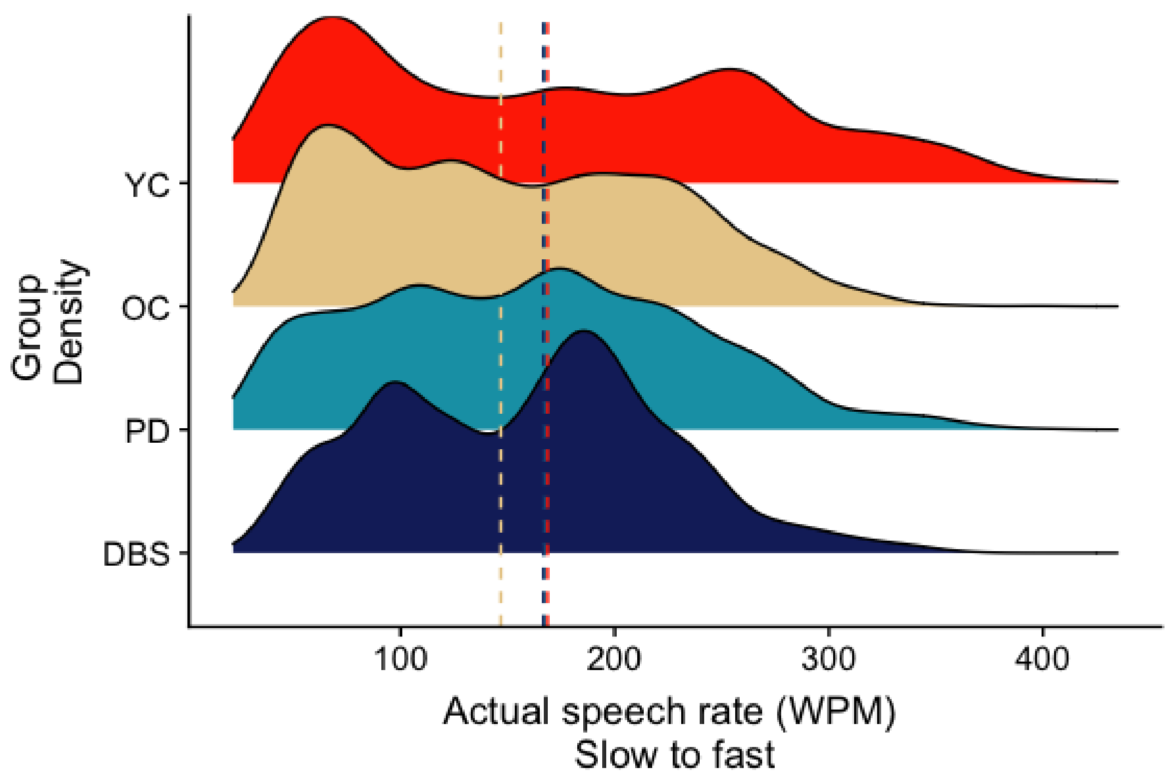


Figure 3: Density plots of the distribution of actual speech rates across groups. Rate was calculated as words per minute for each speaker. Vertical lines indicate mean habitual speech rates for each group.

The distributions of actual speech rate across experimental speech rate conditions appear in Figure 4 A, and proportional speech rate across the final rate bins used in all subsequent analyses in Figure 4 B. Evident in these two figures is the considerable overlap in actual speech rates across the experimental conditions. While, overall, all groups were able to successfully slow and quicken their speech rates as instructed, the extent of this change differed across groups.

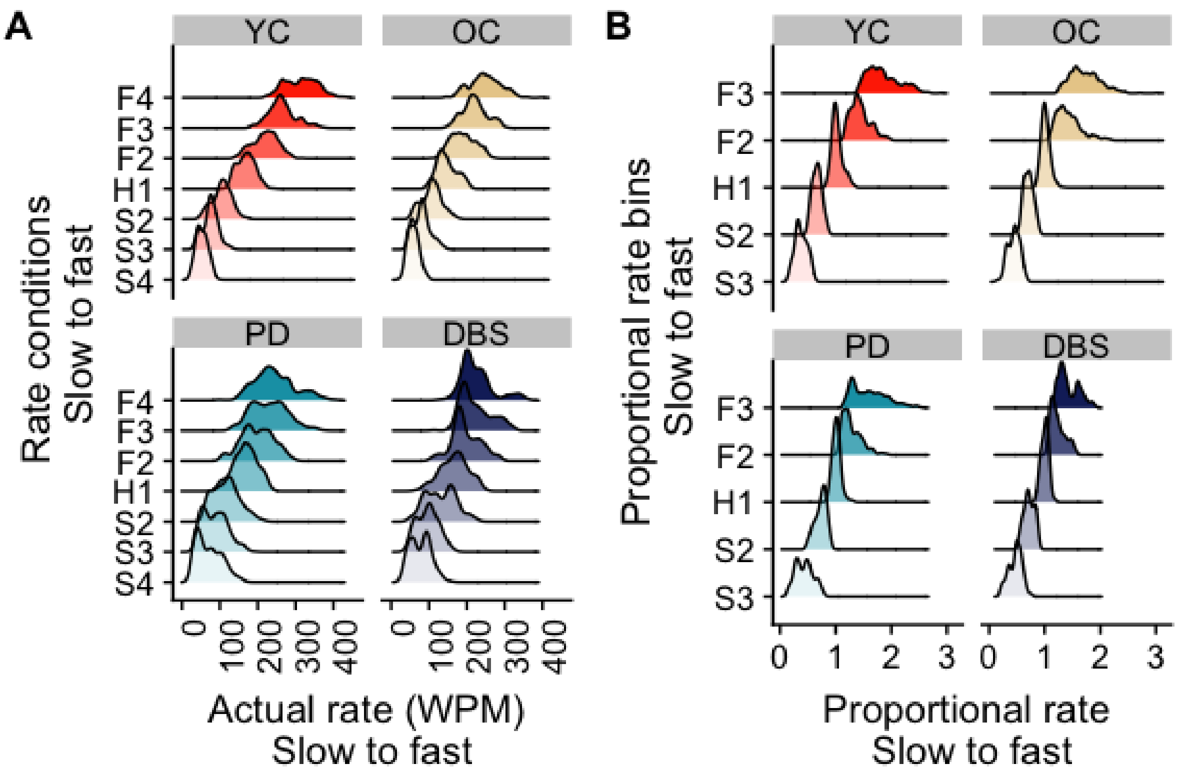


Figure 4: Density plots of the distribution of rates across the categorical rate bins for each group. Figure A displays the distribution of actual speech rate (WPM) across the experimental rate conditions (n = 7). Figure B displays the distribution of proportional rate (where 1 = mean habitual rate) across the final rate bins (n = 5).

4.1.1.2 Speech rate ranges

To explore RQ1 (*to what extent did the groups modify their speech rate?*), slow and fast ranges were calculated for each participant based on actual WPM produced, and a series of Welch Two Sample t-tests were run to test for differences in both *slow* and *fast* speech rate ranges. Average slow and fast ranges for each group are reported in Table 6. Two DBS participants were excluded from this analysis (502, 504) because they did not complete the extreme rate conditions (S4, F4) due to fatigue.

The YC group had the largest range for both the slow and fast rates. For slow speech, the only significant group comparison was between the YC and OC groups, indicating that the younger adults produced a slower range (overall slower speech) than the older healthy adults (YC vs. OC: $p = 0.004$). The comparison between the YC and PD groups approached significance (YC vs. PD: $p = 0.077$), but none of the other comparisons indicated a difference in slow rate ranges (YC vs. DBS: $p = 0.222$; OC vs. PD: $p = 0.326$; OC vs. DBS: $p = 0.31$; PD vs. DBS: $p = 0.852$).

For fast speech, the YC group had a significantly wider range (e.g., produced faster speech) than all the other groups (YC vs. OC: $p = 0.01$; YC vs. PD: $p < 0.001$; YC vs. DBS: $p < 0.001$). The OC group produced a significantly wider range of fast speech compared to the DBS ($p = 0.006$) but not the PD group ($p = 0.149$), and the PD and DBS groups did not differ from each other ($p = 0.12$).

These findings suggest that the PD and DBS groups were successful in slowing their speech rates down to a similar extent to younger and older healthy adults. With regards to faster speech, the younger adults produced a wider range than any of the other groups. The PD and OC groups were similar in their fast speech adjustments, as were the PD and DBS groups. The DBS group, however, were not as able to quicken their speech rate to the same extent as the older healthy adults.

Table 6: Slow and fast speech rate ranges (WPM) for each speaker group

Group		Slow		Fast	
		Range	StDev	Range	StDev
YC	17	-122.918	18.728	185.089	49.223
OC	17	-98.052	26.655	138.337	49.890
PD	22	-107.590	33.219	113.380	55.391
DBS	10	-109.784	29.053	86.042	38.525

The distribution of actual speech rate (WPM) by intended speech rate (2x, 3x, 4x faster/slower, as indexed by the speech rate conditions), are plotted in Figure 5.

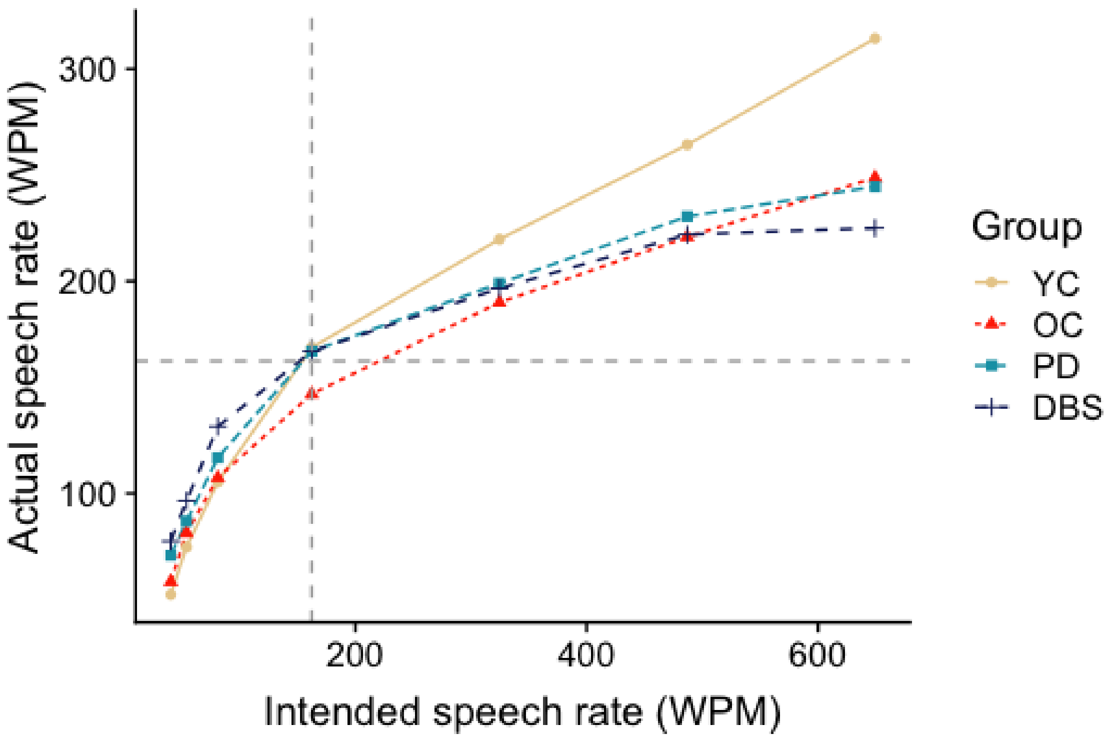


Figure 5: Actual speech rate (WPM; y-axis) as a function of intended speech rate by way of the rate conditions (x-axis). Intended rate based on the grand speech rate mean in habitual rate, indicated by horizontal and vertical dotted lines. Each point represents values for each rate condition, averaged over participants.

4.1.2 Segmental acoustics

Speech rate in this section refers to proportional rate as described in Section 3.2.7. All effects and interactions are reported. Estimates and p-values for null effects are only reported for the main effects (not interactions) for simplicity; all significant estimates and p-values are reported in the text. Post-hoc pairwise comparisons are reported for significant interactions involving both group and rate. The p-values for the post-hoc tests are not directly reported in the text, but appear in the pairwise comparison figures in each section, as well as in Appendix E.

4.1.2.1 Voice onset time

4.1.2.1.1 Final model

There were 311 instances (3.7% of the data) where VOT could not be marked due to the stop being unreleased. For the VOT model, these observations were excluded. The model for (log-transformed) VOT included by-participant random intercepts and slopes for rate (all four contrasts), and by-item random intercepts and slopes for group. All main effects entered into the model improved the model fit and were thus included. Final main fixed effects included group, rate, voicing, PoA, and sex. Two-way interactions that were included were group x rate, group x voicing, voicing x PoA, and voicing x sex. The group x PoA and group x sex, and PoA x sex interactions did not improve the model fit and were thus excluded. Rate interactions with voicing, PoA, and sex also did not improve the fit and were dropped. The only three-way interaction (of those involving group and rate) that improved the model fit was group x rate x voicing. The three-way interactions involving PoA and sex were excluded.

The final fixed effects model structure can be summarized as follows:

VOT (log-transformed) ~ group + rate + voicing + PoA + sex +
group x rate + group x voicing + voicing x PoA +
voicing x sex + rate x voicing + group x rate x voicing

4.1.2.1.2 Main effects

Figures 6 - 12 report the overall trends of VOT. Coefficients are reported in Table 10 in Appendix E.

Speaker group: A main effect of group indicated that VOTs were shortest for the OC group (captured by a marginally significant effect of the OC vs. PD and DBS contrast: $\hat{\beta}$ = -0.129, p = 0.059), longest for the DBS group, and intermediate for the PD group (captured by the PD vs. DBS contrast: $\hat{\beta}$ = -0.192, p = 0.017).

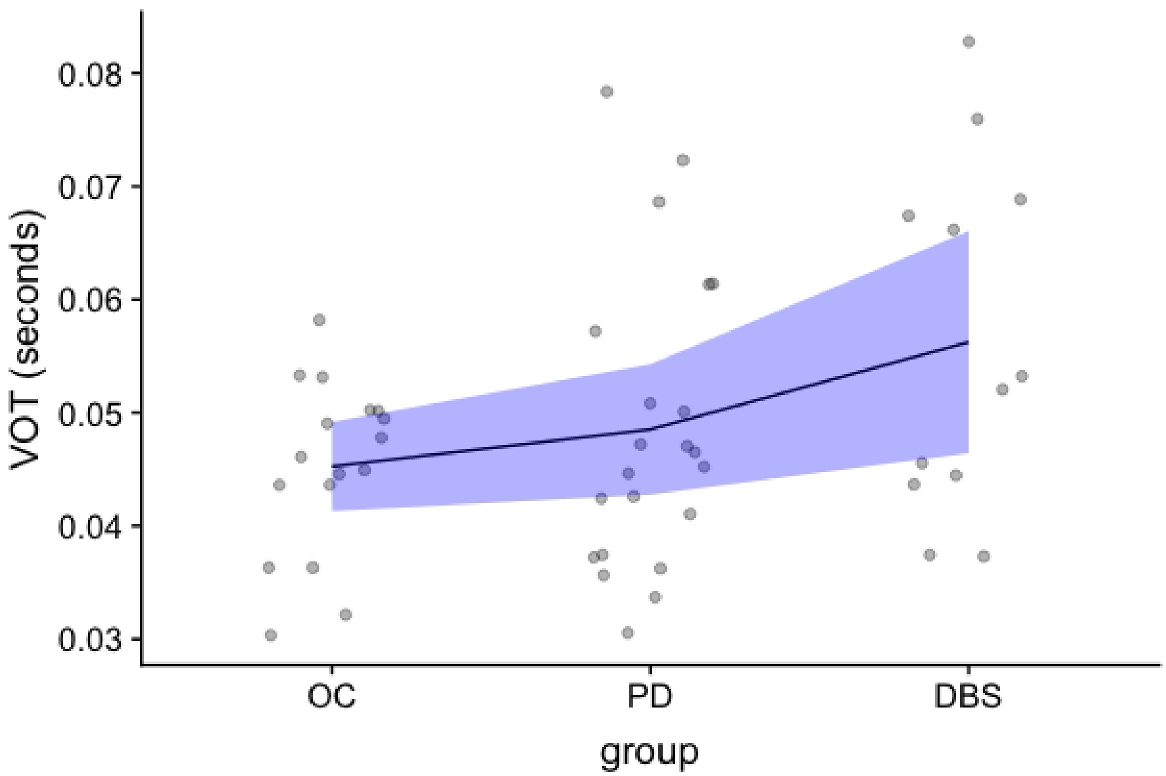


Figure 6: VOT for each speaker group, averaged across participants. Shaded band represents the 95% confidence interval. Points represent average participant values.

Speech rate: The four modified rates (slower = S2; slowest = S3; faster = F2; fastest = F3) were compared with habitual (H1). A significant main effect of any of the modified rates would indicate that that rate was associated with a difference in VOT compared to the habitual rate. Comparisons *across* each of the modified rates are reported in the pairwise comparisons in Table 11 in Appendix E.

All modified speech rates were associated with significant changes in VOT values compared to speakers' habitual rates according to expected trends. Specifically, both slower rates were associated with significantly *longer* VOTs, as indicated by positive estimates (S3 vs. H1: $\hat{\beta}$ = 0.309, p <0.001; S2 vs. H1: $\hat{\beta}$ = 0.223, p <0.001), and both faster rates were associated with significantly *shorter* VOTs (F2 vs H1: $\hat{\beta}$ = -0.174, p <0.001; F3 vs H1: $\hat{\beta}$ = -0.337, p <0.001). This is reported in Figure 7.

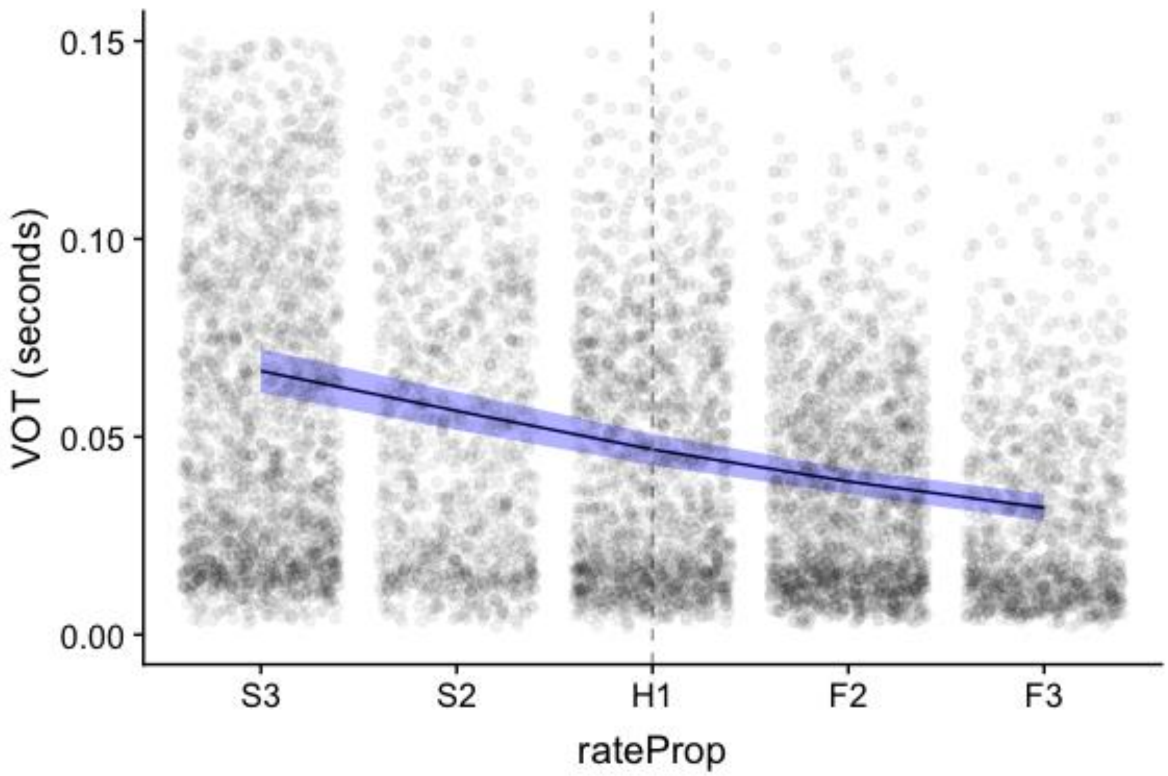


Figure 7: VOT across all speech rates, averaged across participants. Shaded band represents the 95% confidence interval. Points represent individual observations.

Voicing: Voiced stops were associated with significantly shorter VOTs compared to voiceless stops, as would be expected ($\hat{\beta}$ = -0.684, p <0.001). This pattern is visible in Figure 8.

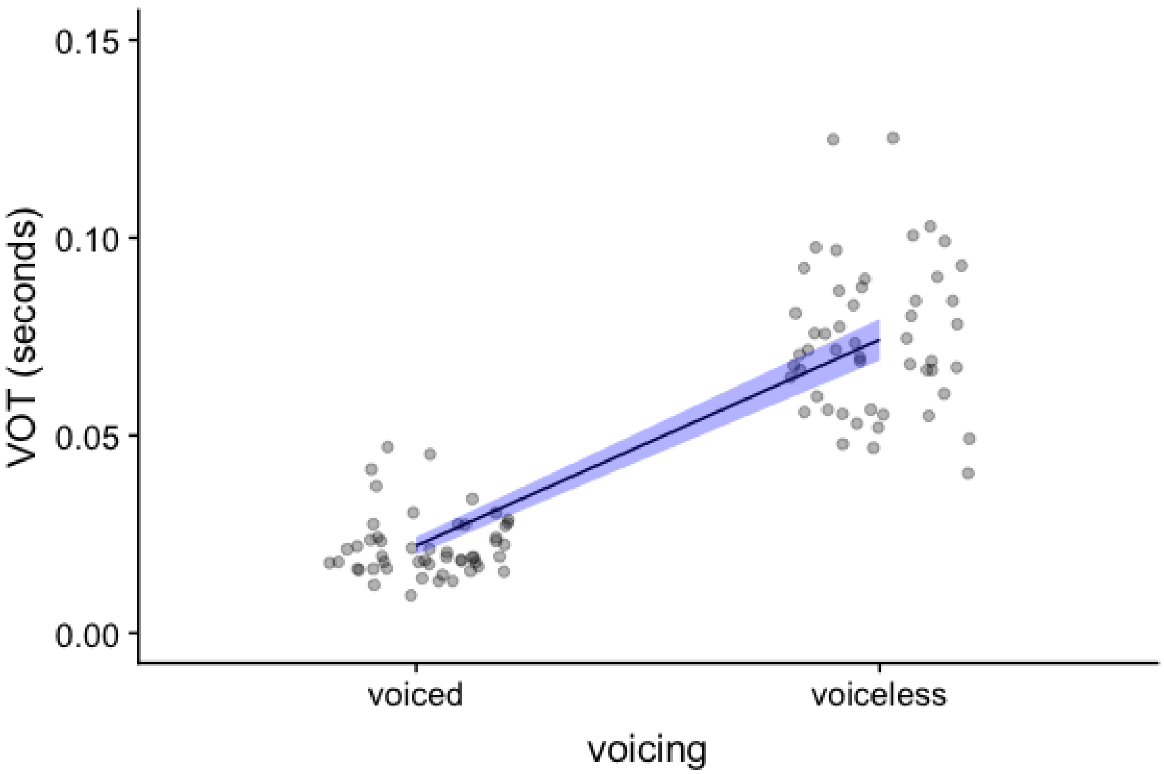


Figure 8: Mean VOT by voicing category, averaged over participants. Shaded band represents the 95% confidence interval. Points represent participant averages.

Place of articulation: VOT was shortest for bilabial stops, followed by alveolar and then velar stops, as seen in Figure 9. This difference was in the expected direction and was significant for both contrast levels (bilabial vs. non-labial: $\hat{\beta}$ = -0.4, p <0.001; alveolar vs. velar: $\hat{\beta}$ = -0.246, p <0.001).

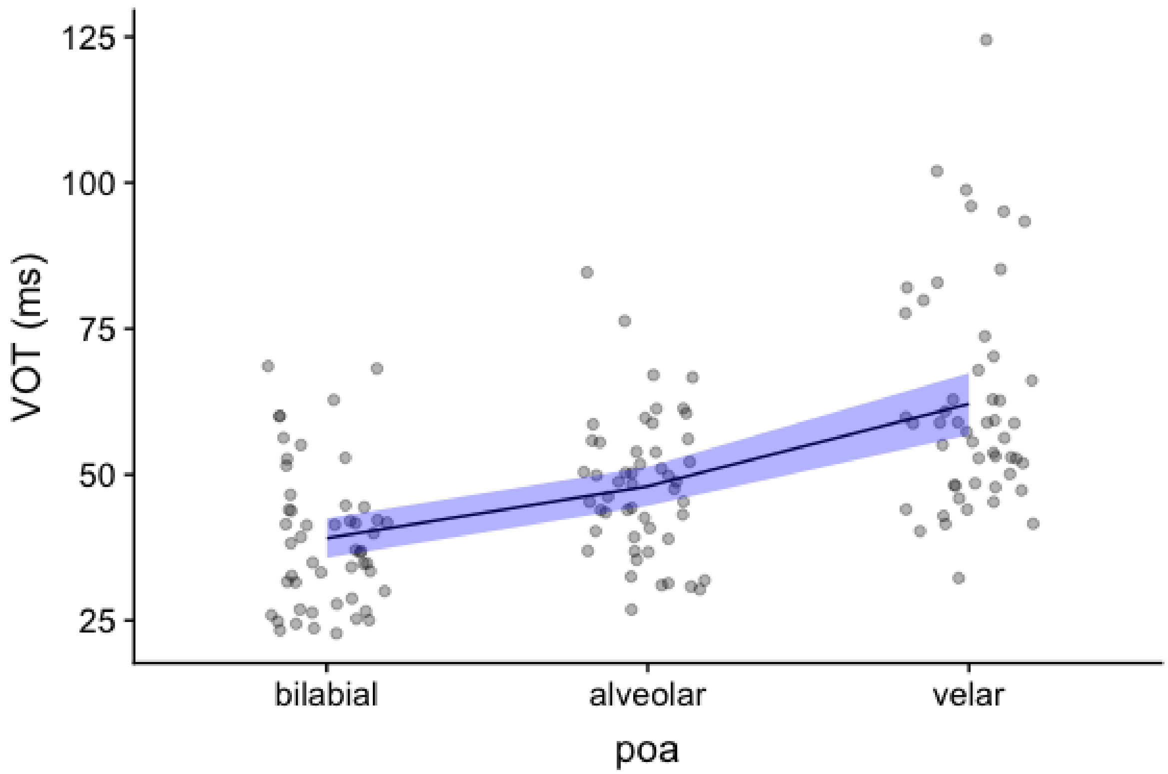


Figure 9: Mean VOT by place of articulation, averaged across participants. Shaded band represents the 95% confidence interval. Points represent participant averages.

Sex: Figure 10 demonstrates that female speakers produced significantly longer VOTs overall compared to male speakers ($\hat{\beta}$ = 0.074, p = 0.033).

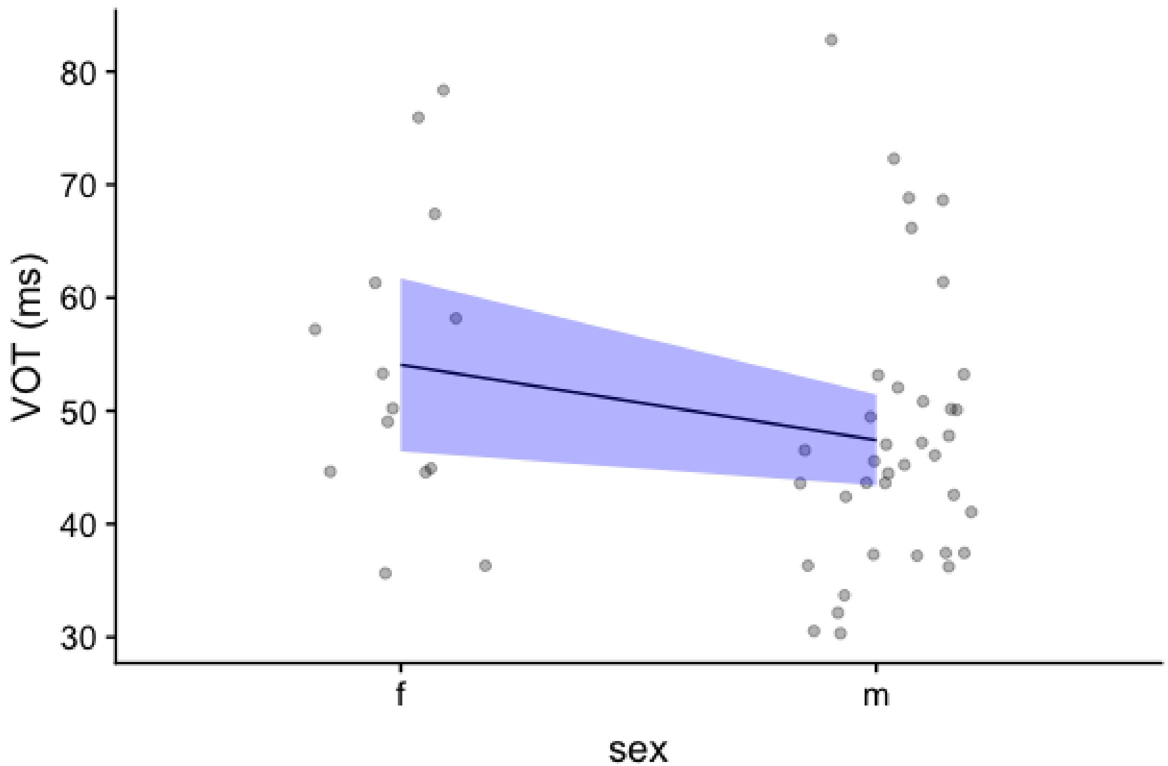


Figure 10: Mean VOT by speaker sex, averaged across participants. Shaded band represents the 95% confidence interval. Points represent participant averages.

4.1.2.1.3 Interactions

Interactions reported here correspond to comparisons for each of the contrast levels specified in the main model. Two sets of post-hoc pairwise comparisons are presented here as well: specifically, within-group differences across speech rates are reported in Figure 13, and within-rate differences across groups appear in Figure 14. The corresponding tables of values used to make these figures appear in Table 11 and Table 12 in Appendix E. All p-values are adjusted using the Tukey HSD method for the number of estimates being compared.

Group by rate interactions: None of the two-way interactions between rate and group were significant. That is, compared to the habitual rate, the groups did not differ significantly in how they modified VOT across slower and faster rates (all else being equal). Note however, that a three-way interaction between group, rate, and voicing did occur, as described below.

Group by voicing interactions: A significant interaction existed for the OC vs. PD and DBS interaction with voicing ($\hat{\beta}$ = -0.045, p = 0.038). Follow-up pairwise comparisons of this two-way interaction demonstrated that the OC group produced shorter voiced and voiceless VOTs than the DBS group, but that this difference was greatest for voiced stops. The PD group, on the other hand, did not significantly differ from either the DBS or OC groups.

Voicing by PoA interactions: Significant interactions were found for both PoA contrast interactions with voicing, indicating that differences between voiced and voiceless VOTs increased with more posterior articulatory positions (bilabial vs. non-labial: $\hat{\beta}$ = -0.079, p <0.001; alveolar vs. velar: $\hat{\beta}$ = -0.07, p = 0.002).

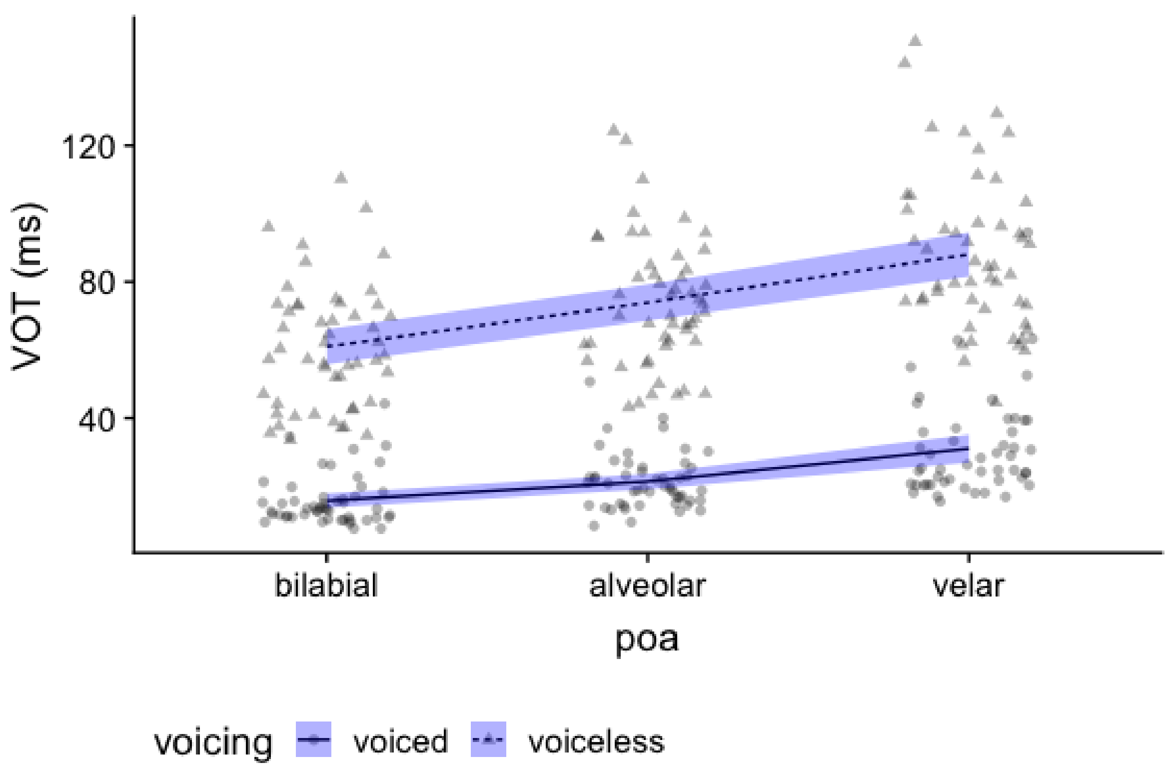


Figure 11: VOT by voicing and place of articulation, averaged across participants. Shaded band represents the 95% confidence interval. Points represent participant averages.

Voicing by speaker sex: A significant interaction between consonant voicing and sex indicated that females demonstrated a larger difference between voiced and voiceless VOTs compared to male speakers ($\hat{\beta}$ = -0.052, p <0.001).

Group, rate, and voicing:

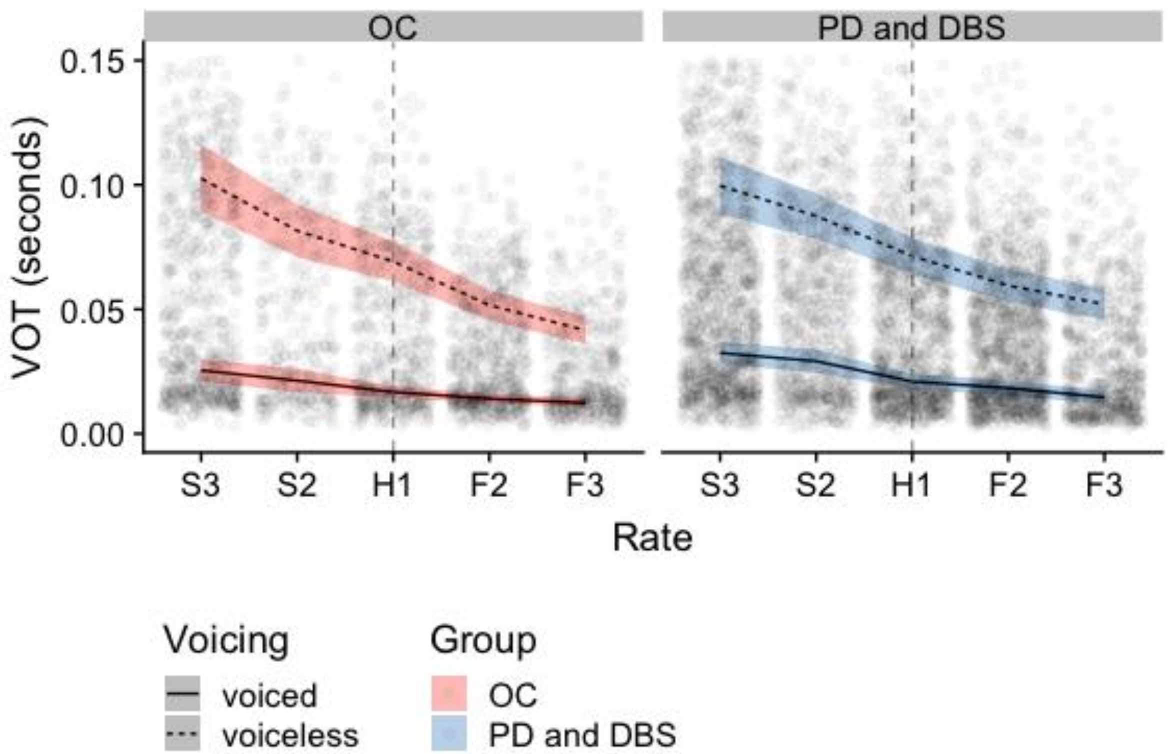


Figure 12: VOT by speaker group (PD and DBS combined), rate, and voicing, averaged across participants. Shaded band represents the 95% confidence interval. Points represent individual observations.

The three-way interaction between group, rate, and voicing is shown in Figure 12. In the fast speech, the OC group produced shorter voiceless VOTs than the PD and DBS groups, and this was significant for both the *faster* and *fastest* rates (compared to habitual) (OC vs. Rest x F2 vs. H1: $\hat{\beta}$ = 0.079, p = 0.009; OC vs. Rest x F3 vs. H1: $\hat{\beta}$ = 0.133, p <0.001). Relatedly, the OC group also produced *longer* voiceless VOT in the *slowest* (S3) speech ($\hat{\beta}$ = -0.059, p = 0.045). No other contrasts for the three-way interaction were significant, including any interactions involving the PD vs. DBS contrasts. In essence, the OC group produced *more contrast* in slowest speech and *less contrast* in fast speech compared to the PD and DBS groups, and this was achieved by larger relative changes to voiceless VOT.

4.1.2.1.4 Pairwise comparisons

Pairwise comparisons are visualized in Figures 13 and 14. Figure 13 reports pairwise comparisons across each proportional rate within each group, and Figure 14 reports pairwise comparisons across each group within each rate. Table 11 and Table 12 in Appendix E report pairwise comparisons for the three-way interaction between group, rate, and voicing.

A note on interpreting the pairwise comparison tables:

Figures like 13 and 14 will appear frequently throughout this Results chapter and are intended to facilitate the reporting of the pairwise comparison results. Numerical values themselves are not reported in the text, but tables corresponding to the pairwise comparison figures are included in the appropriate appendices.

In these figures, the x-axis corresponds to the estimated mean difference between contrasts, and the y-axis denotes the particular contrast levels of interest. Relative effect sizes, then, may be interpreted as the size of the bar for a given contrast (relative to the other effects). Significance levels (p-values) are denoted by color. The meaning of a positive versus negative effect depends on the model structure, and this is always clarified in the text. For the VOT models, for example, a positive comparison indicates that the contrast level specified first was associated with a *larger* (log) VOT than the contrast level specified on the right (so a contrast such as “F2 - F3” with a positive effect may be interpreted as “the faster F2 speech was associated with longer VOT compared to the fastest F3 speech, everything else being equal”).

A negative value in Figures 13 and 14 indicates that the contrast specified on the left was associated with smaller VOT than the contrast on the right. The y-axis represents the primary contrast, and the x-axis represents the pairwise estimated marginal means (Lenth, 2018). In these figures and the ones like them to follow, a significant pairwise comparison is denoted as a red ($p < 0.001$), orange ($p < 0.01$), or yellow ($p < 0.05$) bar.

Regarding the rate comparisons, slower speech was associated with longer VOT for all groups, and this pattern held for both voiced and voiceless stops. As can be seen in Figure 13, most pairwise comparisons between each rate were significant across groups (indicated by red, orange, or yellow bars). Some exceptions were as follows. VOT values in the two slowest rates (S3 vs. S2) did not significantly differ for either group for either voiced or voiceless stops. VOT values in the two fastest rates (F2 vs. F3) did not differ for the PD's voiced stops, nor for the DBS group's voiceless stops.

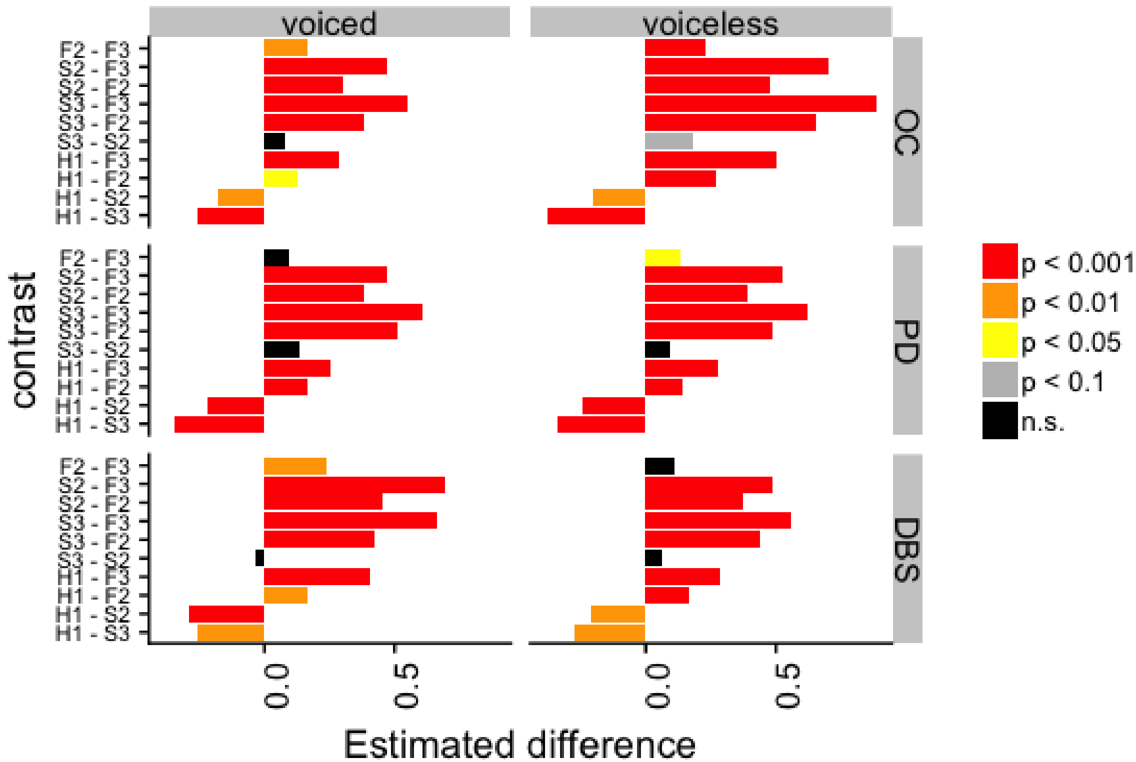


Figure 13: VOT pairwise comparisons for the three-way interaction between group, rate, and voicing, illustrating differences between proportional rates. X-axis represents estimated difference of the mean for the model. Response variable is on a log scale. P-values were adjusted using the Tukey HSD method.

Figure 14 reports the pairwise comparisons between groups across each rate for voiced and voiceless stops. Overall, more comparisons reached significance for voiced stops. In particular, within the habitual rates, both the OC and PD groups demonstrated smaller voiced VOTs than the DBS group, though the difference between the OC and PD groups was not significant. There were no group differences for voiceless VOTs, though the same directional pattern was observed.

Voiceless VOTs did not differ for either group in slower speech. In the slowest rate (S3), voiced VOTs were shorter for the OC group compared to the DBS group. The other group comparisons were not significant. In the slower rate (S2), both OC and PD groups demonstrated smaller voiced VOTs compared to the DBS group and did not differ from each other.

In the faster rate (F2), once again, the OC and PD groups produced smaller voiced VOTs compared to the DBS group but did not differ from each other. The OC group also produced significantly smaller voiceless VOTs compared to the DBS group at this rate.

No significant differences between groups was observed for voiced stops at the fastest rate (F3). The OC group produced significantly smaller voiceless VOTs than both the PD and DBS groups at this rate.

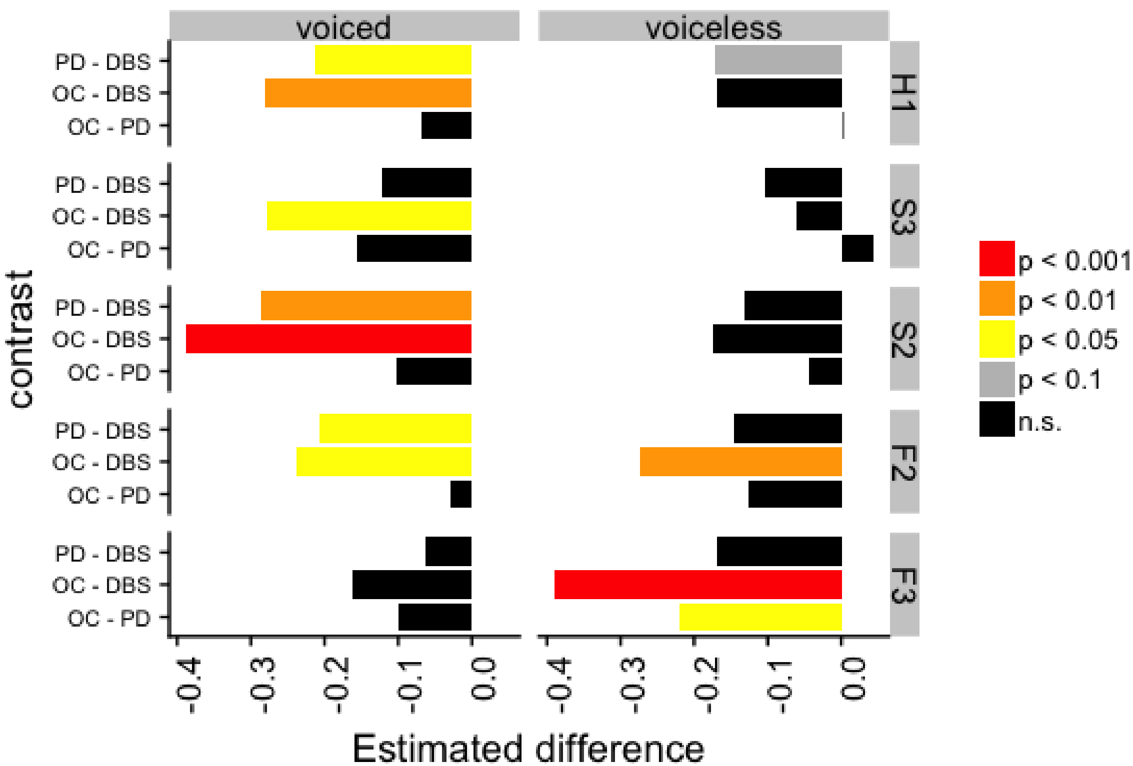


Figure 14: VOT pairwise comparisons for the three-way interaction between group, rate, and voicing, illustrating differences between groups. X-axis represents estimated difference of the mean for the model. Response variable is on a log scale. P-values were adjusted using the Tukey method.

4.1.2.2 Voicing during closure

4.1.2.2.1 Final model

Voicing during closure could not be measured on 208 consonants (2.5%) either due to closure being unidentifiable (typically due to complete spirantization; n = 172) or closure being too short for the pitch tracker to detect voicing (i.e., < 20 ms; n = 36). These tokens were removed from the analysis.

As reported in the methods, VDC was treated dichotomized into two categories: total voicing and some or no voicing through closure. VDC was then modelled using logistic mixed-effects regression using the glmer() function of the lme4 R package (Bates et al., 2015).

The random effects structure in the final model included random by-participant random intercepts and slopes for all rate contrasts, and by-item random intercepts. By-item random slopes for group led to a singular fit and were thus dropped.

Fixed effects in the final model included group, rate, voicing, and PoA. Sex did not improve the model fit and was dropped. Two-way interactions included group x rate, rate x voicing, group x voicing, rate x PoA, and voicing x sex. Most two-way interactions involving PoA and sex did not improve the model fit and were dropped. None of the three-way interactions significantly improved the model fit.

The final fixed effects structure for the VDC model can be summarized as follows:

VDC (binary) ~ group + rate + voicing + PoA + group x rate +
group x voicing + rate x voicing + rate x PoA + voicing x sex

4.1.2.2.2 Main effects

Figures 15 - 18 report the overall trends for VDC. Note that these figures include VDC as a continuous variable from 0 to 1, but the model included a dichotomized VDC (total versus some or none). The treatment of VDC as continuous in the figures is to better visualize the overall patterns in the data. Coefficients are reported in Table 13 in Appendix E. Positive model coefficients may be interpreted as a greater occurrence of stops with voicing occurring through the entire closure (e.g., VDC = 1).

Speaker group: There was no main effect of speaker group for either contrast (OC vs. PD, DBS: $\hat{\beta}$ = -1.104, p = 0.158; PD vs. DBS: $\hat{\beta}$ = -0.114, p = 0.9).

Speech rate: Compared to habitual speech rates, the slowest rate (S3) was associated with significantly less voicing through stop closure ($\hat{\beta}$ = -1.046, p = 0.005), the slightly slower rate (S2) did not significantly differ ($\hat{\beta}$ = -0.553, p = 0.117), and both faster rates were associated with significantly more voicing through closure (F2: $\hat{\beta}$ = 1.464, p <0.001; F3: $\hat{\beta}$ = 2.73, p <0.001). This pattern can be seen in Figure 15.

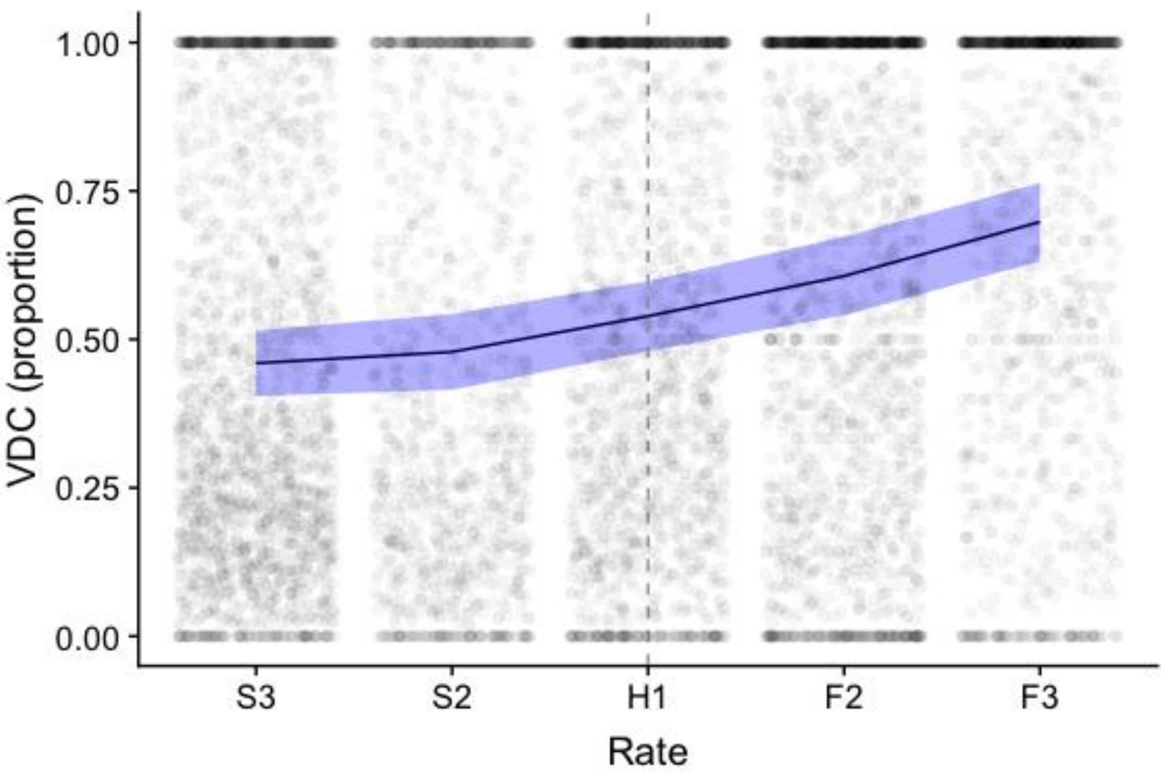


Figure 15: Proportion of voicing during closure by speech rate, averaged across participants. Shaded band represents the 95% confidence interval. Points represent individual observations.

Voicing: As expected, a greater proportion of VDC was significantly associated with voiced stop consonants ($\hat{\beta}$ = 3.112, p <0.001).

Place of articulation: VDC showed a tendency to increase with more posterior places of articulation. Bilabial stops were associated with less VDC compared to alveolars and velars ($\hat{\beta}$ = -0.562, p = 0.006), and alveolars had less VDC than velars ($\hat{\beta}$ = -0.874, p <0.001).

4.1.2.2.3 Interactions

Neither the two-way interaction between group and rate, nor the three-way interactions involving group, rate, and voicing or PoA were significant. As such, no post-hoc comparisons are reported for this measure. All two-way interactions are reported here.

Group by rate interactions: None of the two-way interactions involving rate (comparing the modified rates with habitual) and speaker group were significant (8 interactions in total: 4 rate contrasts by 2 group contrasts). In other words, all three groups demonstrated similar amounts of VDC at each rate.

Group by voicing interactions: Both group comparisons demonstrated a significant interaction with stop voicing, indicating that the OC group demonstrated more voicing through closure of voiceless stops compared to the PD and DBS groups (OC vs. PD, DBS: $\hat{\beta}$ = 0.488, p = 0.006), and the PD group demonstrated more than the DBS group (PD vs. DBS: $\hat{\beta}$ = 0.62, p <0.001).

Figure 16 A reports the distribution of VDC across voicing contrasts for each group. The bottom panel of this plot (voiceless stops) demonstrates a sharper peak for the OC group in the direction of little to no voicing, whereas the distribution of VDC progressively flattens for the PD and DBS groups. This may indicate that the stop voicing contrast was better maintained (at least for this cue) for healthy talkers, who produced less VDC in voiceless stops.

Rate by voicing interactions: The interaction between the fastest rate (F3) and voicing was the only significant comparison, indicating that the difference in VDC between voiced and voiceless stops was most similar at this fastest rate when compared to habitual ($\hat{\beta}$ = -0.908, p <0.001).

Figure 16 B reports the distribution of VDC for the voicing contrast across all speech rates. Figure 17 demonstrates this pattern for each of the three groups. While subtle, this figure suggests a steeper rise towards more VDC in the fast speech for voiceless stops compared to voiced stops, accounting for a greater approximation between these categories at the fastest rate. While there was no three-way interaction including group, visually this trend is most apparent in the empirical data for the DBS group.

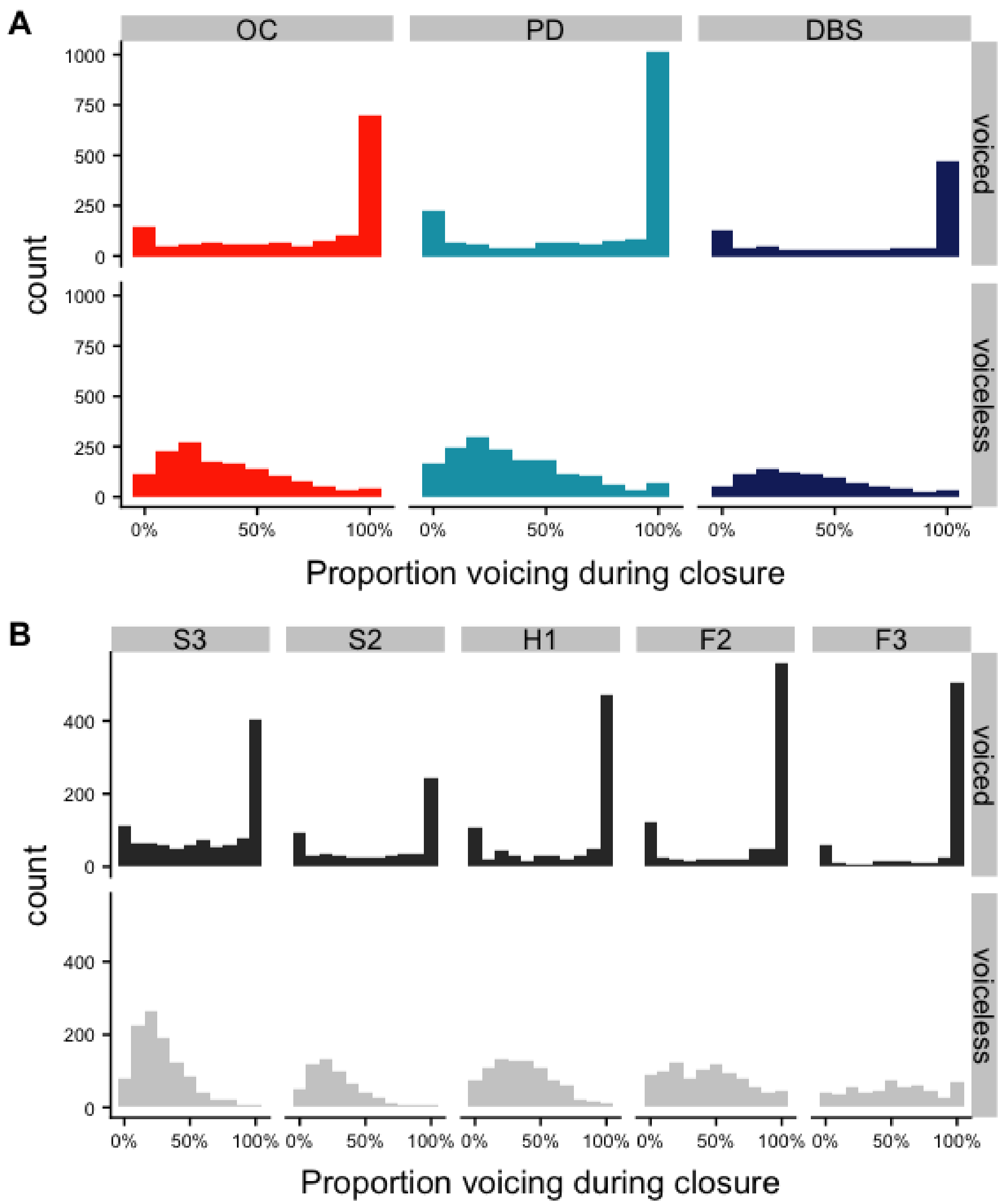


Figure 16: Distribution of the proportion of voicing during closure. Figure A is a density histogram displaying the proportion of VDC for voiced and voiceless stops by speaker group. Figure B is a histogram (bin width = 0.1) displaying the proportion of VDC across speech rates for voiced and voiceless stops.

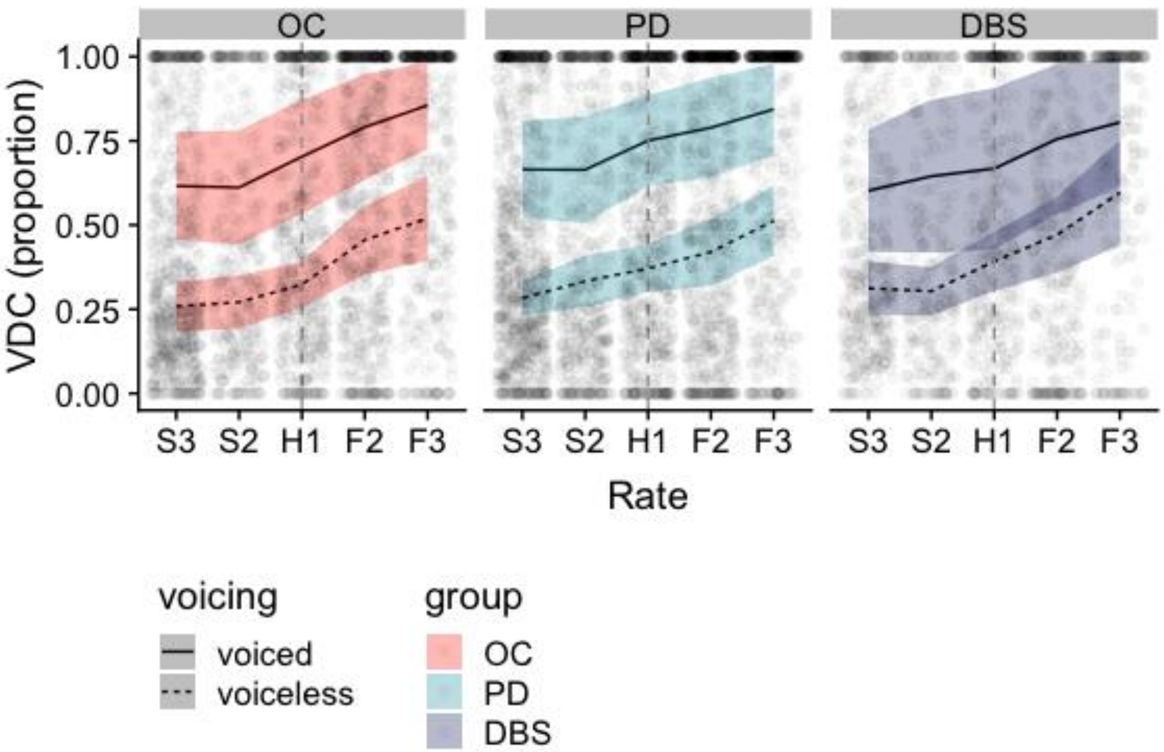


Figure 17: Proportion of voicing during closure by speech rate, speaker group, and stop voicing status, averaged across participants. Shaded band represents the 95% confidence interval. Points represent individual observations.

Rate by PoA interactions: The interaction between rate and the bilabial vs. non-bilabial contrast was only significant for the slowest rate (S3: $\hat{\beta}$ = 0.572, p = 0.035). Conversely, the rate interaction with the alveolar vs. velar contrast was only significant for the fastest rate (F3: $\hat{\beta}$ = 1.093, p = 0.002). In general, velar stops were associated with more VDC compared to bilabial and alveolar stops. Figure 18 reports the empirical data across rate, PoA, group, and voicing. While all three-way interactions were excluded from the final model because they did not significantly improve the model fit, a four-way display of the relationship between group, rate, voicing, and PoA appears in Figure 18 to facilitate visualization of the patterns described above.

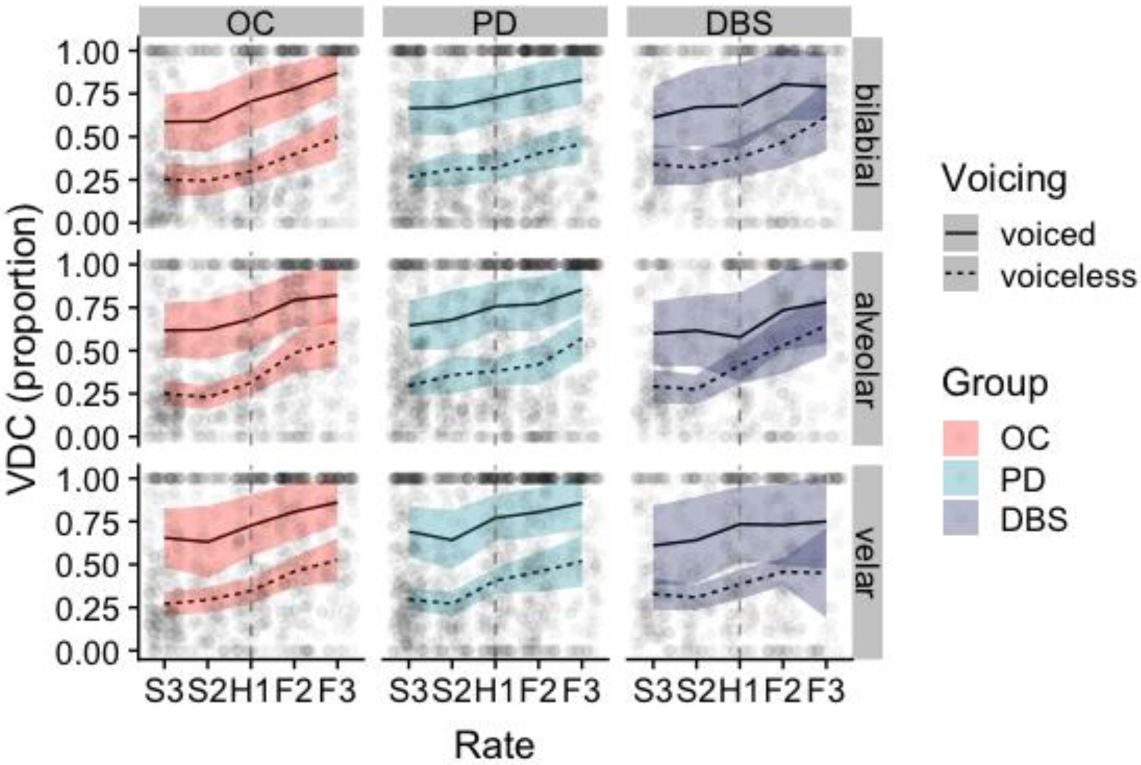


Figure 18: Voicing during closure by speaker group, speech rate, place of articulation, and voicing, averaged across participants. Shaded band represents the 95% confidence interval. Points represent individual observations.

4.1.2.3 Quadrilateral vowel articulation index

Quadrilateral vowel index (QVAI) was modelled as a function of group, rate, and speaker sex, as well as their interactions. A larger QVAI value may be interpreted as less vowel centralization, i.e., a larger vowel space. To reiterate, because the QVAI metric must, by definition, be calculated as an index of multiple vowels (specifically four), it was not possible to run the models on the raw data, in which each observation included the production of one nonce word (containing a unique consonant vowel pair). A summary data set was thus calculated in the following way. For each level of the proportional rate to be included in the model (five levels in total: H1, S2, S3, F2, F3), the formant values for each vowel (/i, æ, ɑ, u/) were averaged for each speaker. This summary data set was constructed using the same raw data from the previous models, so, the utterances containing stop consonants.

Two speakers in the control group (OC208 and OC212) did not have data in the S2 bin, due to the extent to which they slowed their speech (i.e., they did not produce intermediate values between their slowest and habitual rates as captured by the proportional binning procedure). Thus, two data points in this summary data were omitted, resulting in 253 observations in total (51 speakers x 5 rate bins, minus 2 data points).

Only three variables of interest were thus considered: group, rate, and speaker sex. As with the VOT and VDC models, the model started with group, sex, and their interaction. The addition of sex as a fixed effect, as well as the interaction between rate and sex improved the model fit. The two- and three-way interactions involving group and sex did not improve the model fit and thus were not included. While the other models included by-speaker and by-item random effects terms, the QVAI model only included by-speaker random effects, as individual item information was collapsed in the summary data. The random effects structure included by-participant random intercepts and slopes for rate (all four contrasts).

The final fixed effects structure for the QVAI model can be summarized as:

$$\text{QVAI} \sim \text{group} + \text{rate} + \text{sex} + \text{group } x \text{ rate} + \text{rate } x \text{ sex}$$

4.1.2.3.1 Main effects

Figures 19 - 22 report the overall trends of QVAI. A larger QVAI value indicates greater vowel space. The group x rate interaction was not significant, and as such, no pairwise comparisons are reported.

Speaker group: QVAI was largest for the OC group and smallest for the DBS group, indicated by significant differences for both group contrasts (OC vs. rest: $\hat{\beta}$ = 0.076, p = 0.043; PD vs. DBS: $\hat{\beta}$ = 0.088, p = 0.045). Figure 19 demonstrates this trend, and Figure 21 A displays the vowel quadrilaterals in F1 and F2 space for all groups.

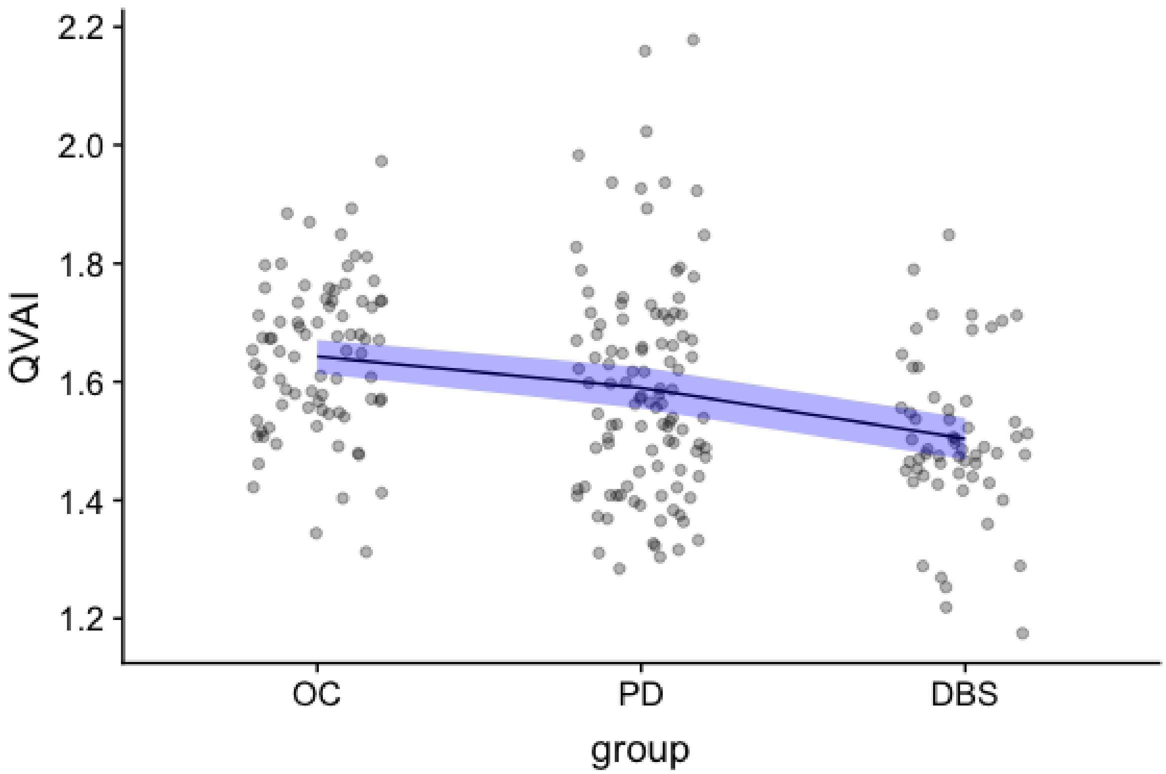


Figure 19: QVAI by speaker group. Points represent individual observations for each participant and rate. Shaded band represents the 95% confidence interval.

Speech rate: Figure 20 reports QVAI as a function of rate, and Figure 21 B reports the distribution of QVAI for each group at each speech rate. When other variables were held constant, vowel centralization was evident in both faster rates compared to habitual speech (F2 vs. H1: $\hat{\beta}$ = -0.077, p <0.001; F3 vs. H1: $\hat{\beta}$ = -0.128, p <0.001). This is visible as a downward slope in the fast conditions in Figure 20 and as a leftward shift in peaks towards a smaller QVAI value in the bottom panels in Figure 21. While there was a trend for QVAI to increase in the slower speech, this comparison was only marginally significant for the slowest rate (S3 vs. H1: $\hat{\beta}$ = 0.038, p = 0.051; S2 vs. H1: $\hat{\beta}$ = 0.022, p = 0.118). In other words, vowel space, as represented by the vowel articulation index, trended towards expansion in slower speech, and became significantly smaller (centralized) in faster speech. As can be seen in Figures 20 and 21 B, QVAI is overall larger in slower rates, but also more variable.

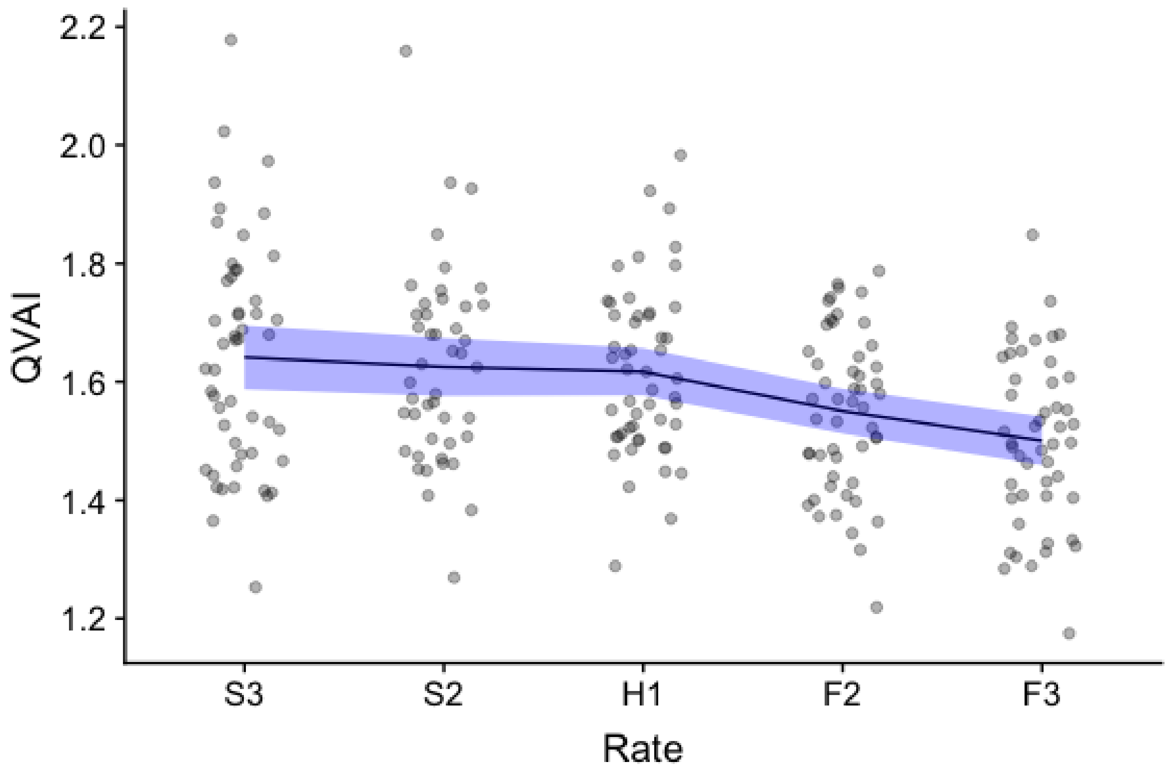


Figure 20: QVAI by speech rate. Points represent individual observations. Shaded band represents the 95% confidence interval.

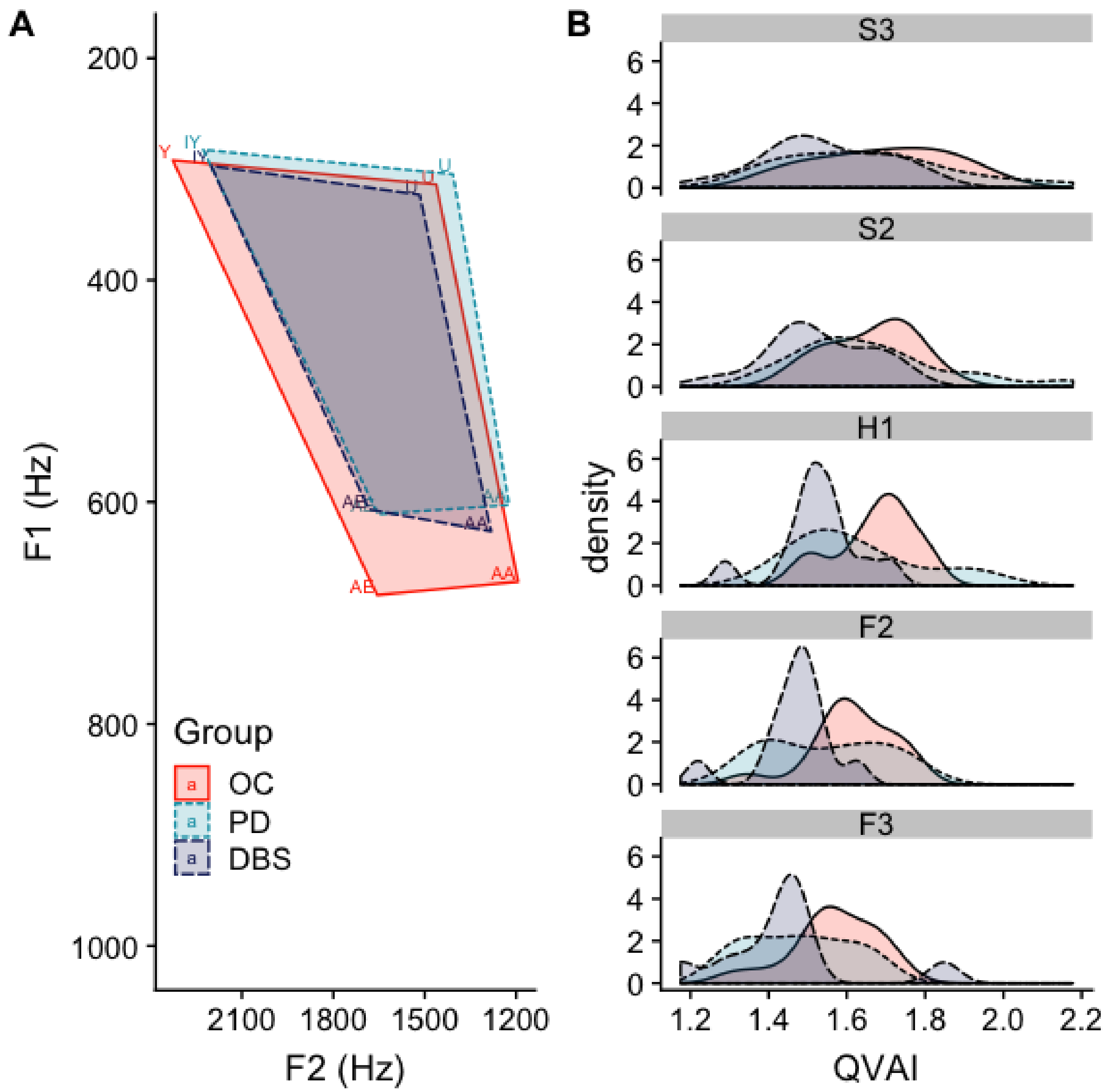


Figure 21: Figure A displays the vowel polygons produced in the first and second formant space by each speaker group. Figure B is a set of density plots showing the distribution of QVAI for each group at each rate (ordered top to bottom from slow to fast).

Sex: Female speakers produced significantly larger vowel spaces compared to male speakers ($\hat{\beta}$ = 0.053, p = 0.007).

4.1.2.3.2 Interactions

None of the possible two-way interactions between rate and group were significant for QVAI (Table 14 in Appendix E). Interactions between rate and sex were significant for the faster rates (F2 vs. H1: $\hat{\beta}$ = -0.027, p = 0.019; F3 vs. H1: $\hat{\beta}$ = -0.128, p <0.001), but not for the slower rates. Figure 22 demonstrates that at faster rates, males and females had greater overlap in their degree of vowel centralization.

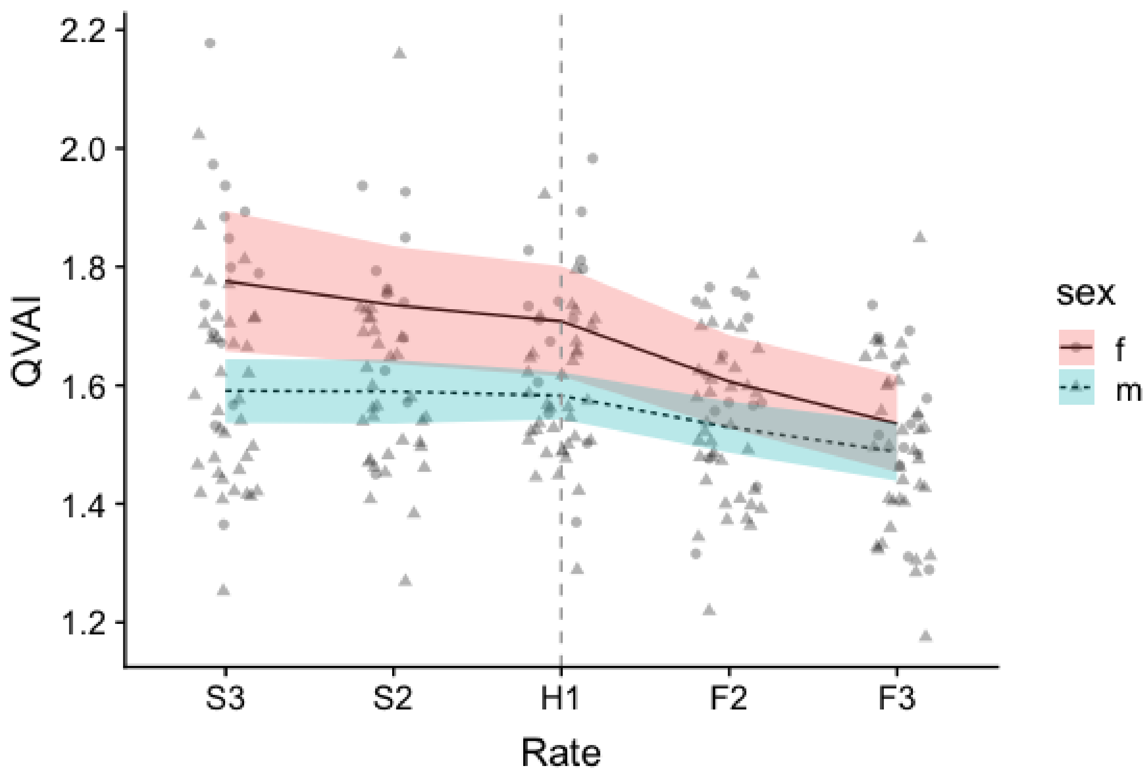


Figure 22: QVAI by speech rate and speaker sex. Points represent individual observations. Shaded band represents the 95% confidence interval.

4.1.3 Voice acoustics

Voice acoustics included speech intensity and harmonics-to-noise ratio (HNR), which were measured on the vowels of interest in the nonce words.

4.1.3.1 Intensity

4.1.3.1.1 Final model

The predictors of interest for intensity included group, rate, and sex. As with the previous models, the base model included group, rate, and their interaction. Including by-item random slopes for speaker group led to a singular model fit and were thus excluded. Random effects terms included by-item random intercepts and by-speaker random intercepts and slopes for rate (all contrasts).

The final model included fixed effects of group, rate, and sex. No two-way interactions involving sex were included, but the three-way interaction was. The final fixed effects structure can be summarized as:

Intensity ~ group + rate + sex + group x rate + group x sex +
rate x sex + group x rate x sex

4.1.3.1.2 Main effects

Figures 23 through 24 report the empirical data for intensity. Model coefficients are reported in Table 15 in Appendix E. Figure 25 reports the pairwise comparisons for the two-way group x rate interaction, which also appear in Table 16 in Appendix E.

Speaker group: There was no main effect of speaker group (OC vs. Rest: $\hat{\beta}$ = 1.471, p = 0.169; PD vs. DBS: $\hat{\beta}$ = 1.275, p = 0.383), indicating that mean intensity (measured on the vowel) did not significantly differ across the OC, PD, and DBS groups.

Speech rate: A main effect of rate was found for the two slow speech contrasts (S3 vs. H1: $\hat{\beta}$ = -2.798, p <0.001; S2 vs. H1: $\hat{\beta}$ = -2.042, p <0.001), indicating that both slower speech conditions were associated with lower speech intensity compared to habitual speech. The fast speech conditions did not significantly differ from habitual (F2 vs. H1: $\hat{\beta}$ = 0.168, p = 0.688; F3 vs. H1: $\hat{\beta}$ = 0.511, p = 0.282).

Sex: Males had higher speech intensity compared to females ($\hat{\beta}$ = -1.365, p = 0.016).

4.1.3.1.3 Interactions

Group by rate interactions

Significant group interactions with the slowest rate (S3) demonstrated that while all groups got quieter in slower speech, the OC group did this to a lesser extent than the clinical groups (OC vs. PD x S3 vs. H1: $\hat{\beta}$ = 1.486, p <0.001), and that this was largely driven by the DBS group, who were quieter than the PDs at the slowest rates (PD vs. DBS x S3 vs. H1: $\hat{\beta}$ = 1.709, p <0.001). The same trend was apparent in the slower (S2) speech, but did not reach significance (OC vs. PD: $\hat{\beta}$ = 0.687, p = 0.07; PD vs. DBS: $\hat{\beta}$ = 0.939, p = 0.062). This can effectively be seen as a steeper downward slope for the DBS group compared to the others as speech rate decreases in Figure 23. Pairwise comparisons are reported in Figure 25.

Group, rate, sex interactions

The three-way interaction demonstrated significant differences between group, rate, and sex for comparisons between *slow* and habitual speech only. Specifically, OC male and female speakers demonstrated similar speech intensity in slow speech, which resulted from the male speakers reducing their speech intensity to a greater extent than the females. Conversely, female and male speakers with PD (including those with DBS) both decreased their speech intensity by similar amounts, such that the male-female difference in intensity was preserved at slower rates. This interaction was significant for both slow speech rates for the OC vs. PD, DBS contrasts (S2 vs. H1: $\hat{\beta}$ = 1.032, p = 0.007; S3 vs. H1: $\hat{\beta}$ = 2.163, p <0.001). While the PD and DBS male and female speakers also differed at the slowest rate (PD vs DBS x S3 vs. H1: $\hat{\beta}$ = 0.835, p = 0.003), however, this interaction should be considered with extreme caution, as the DBS group (n=12) only had two female participants. This interaction can be seen in Figure 24.

4.1.3.1.4 Pairwise comparisons

Pairwise comparisons are reported in Figure 25 and in Table 16 in Appendix E. The trend for lower intensity in slower speech was statistically significant for most pairwise comparisons (Figure 25) across speech rates with the following exceptions. Neither of the fast conditions (F2, F3) significantly differed from one another nor from the habitual conditions for any group. The two slow rates did not differ from one another for the PD group or the OC group, but the DBS group demonstrated lower speech intensity in the S3 vs. S2 rates. Speech intensity between the "slower" and "fastest" speech (S2 vs. F3) did not reach significance for the DBS group (but did for the OC and PD groups).

In summary, when speech rate was modified, slower speech was consistently produced more quietly than faster speech. Habitual speech was also produced at a louder volume than slower speech but did not differ from faster speech.

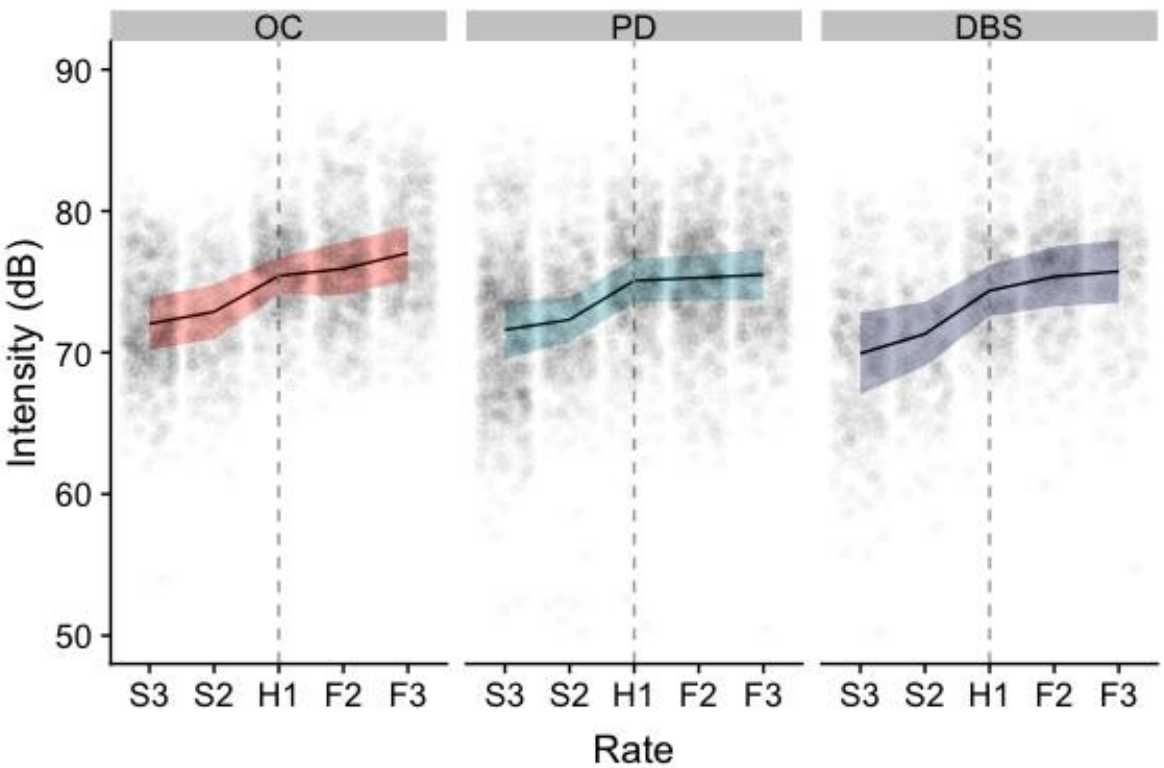


Figure 23: Vowel intensity (dB) by speaker group and speech rate, averaged over participants. Shaded band represents the 95% confidence interval. Points represent individual observations.

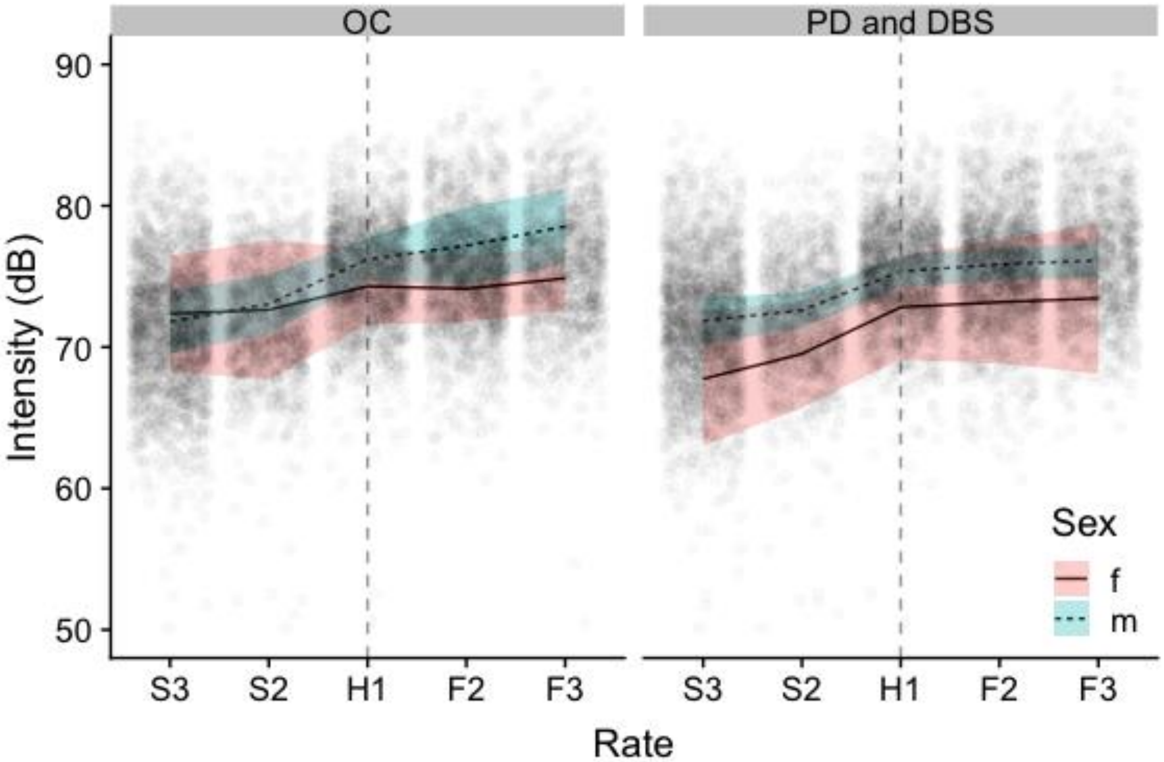


Figure 24: Vowel intensity (dB) by speaker group (PD and DBS groups combined), speech rate, and speaker sex, averaged over participants. Shaded band represents the 95% confidence interval. Points represent individual observations.

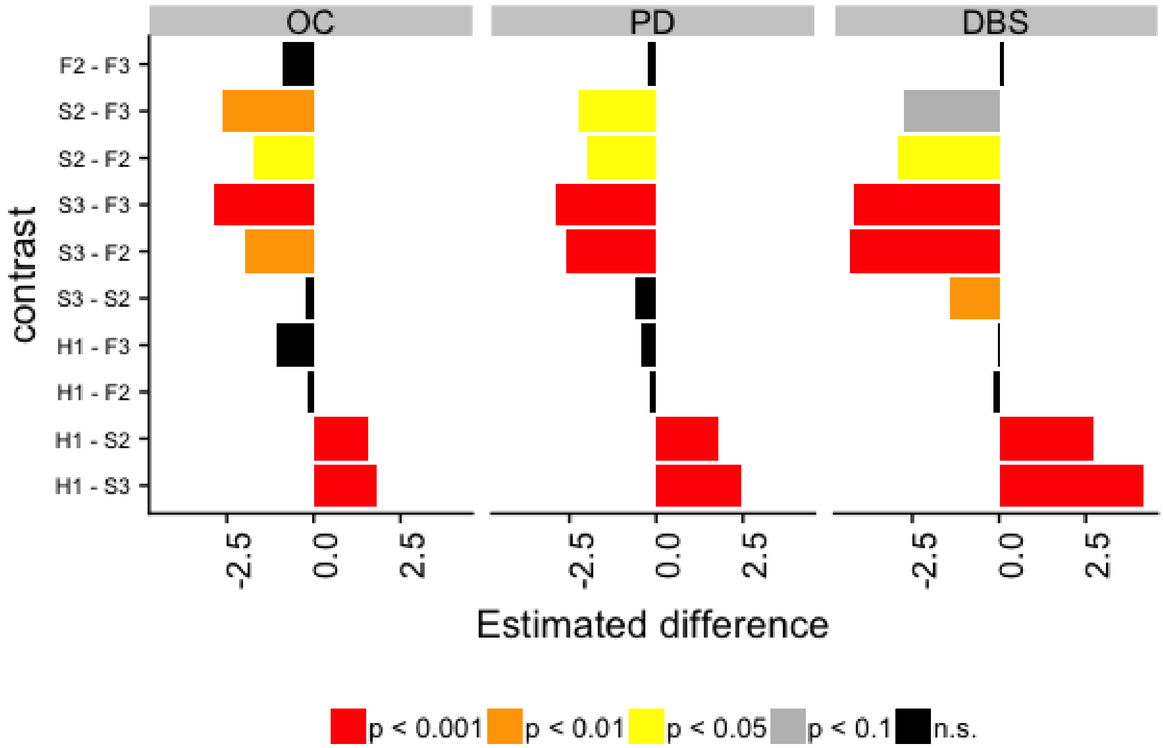


Figure 25: Vowel intensity pairwise comparisons for the two-way interaction between group and rate. X-axis represents estimated difference of the mean for the model. P-values were adjusted using the Tukey HSD method.

4.1.3.2 Harmonics-to-noise ratio

4.1.3.2.1 Final model

HNR was modelled as a function of group, rate, speaker sex, and vowel height. Including by-item random slopes for speaker group led to a singular model fit, and thus these were excluded. Random effects terms in the final model included by-item random intercepts, and by-speaker random intercepts and slopes for rate (all contrasts).

The final model included fixed effects of group, rate, and vowel height. Sex did not improve the model fit. All two-way interactions with the exception of group x sex improved the model fit and were included. Both possible three-way interactions involving group and rate (vowel height, sex) were also included. The fixed effects structure can be summarized as follows:

$$\text{HNR} \sim \text{group} + \text{rate} + \text{vowel height} + \text{sex} + \text{group } x \text{ rate} + \text{group } x \text{ vowel height} + \text{group } x \text{ sex} + \text{rate } x \text{ vowel height} + \text{rate } x \text{ sex} + \text{vowel height } x \text{ sex} + \text{group } x \text{ rate } x \text{ vowel height} + \text{group } x \text{ rate } x \text{ sex}$$

4.1.3.2.2 Main effects

Figures 26 and 28 report the empirical data for HNR. A larger HNR value indicates *better* voice quality. Model coefficients are reported in Table 17 in Appendix E. Pairwise comparisons for the three-way interaction between group, rate, and vowel height are reported in Figure 27 and the values appear in Table 18 in Appendix E.

Speaker group: With all other variables held equal, HNR did not differ significantly by speaker group (OC vs. PD, DBS: $\hat{\beta}$ = 0.231, p = 0.821; PD vs. DBS: $\hat{\beta}$ = -0.569, p = 0.685).

Speech rate: Most modified speech rates (with the exception of S2) were associated with significant *decreases* in HNR compared to habitual speech (S3vH1: $\hat{\beta}$ = -1.106, p = 0.012; F2vH1: $\hat{\beta}$ = -0.657, p = 0.026; F3vH1: $\hat{\beta}$ = -1.324, p <0.001). That is, with all other variables held at their constant values, modifying one's speech rate in either direction was associated with lower HNR (worse voice quality).

Vowel height: High vowels were associated with higher HNR (better voice quality) compared to low vowels ($\hat{\beta}$ = 3.152, p <0.001).

4.1.3.2.3 Interactions

Group by rate interactions

Significant interactions between group and speech rate are apparent in Figure 26. Overall, the OC and PD groups demonstrated better voice quality during slower speech and worse voice quality during faster speech, compared to their habitual rates. The DBS group, on the other hand, displayed the opposite pattern; their voice quality worsened in slower speech and improved in faster speech. This effect was significant in the slower but not faster speech comparisons, with the exception that the PD vs. DBS contrast was not significant for the S2 rate (OC vs. rest in S3: $\hat{\beta}$ = 2.33, p = 0.005; OC vs. rest in S2: $\hat{\beta}$ = 1.412, p = 0.034; PD vs. DBS in S3: $\hat{\beta}$ = 4.042, p <0.001).

Group, rate, and vowel height interactions

These patterns were also apparent across vowel height categories, though to different degrees. High vowels, which were produced with higher HNR overall, were subject to flatter improvements for the OC and PD groups and steeper decrements for the DBS group, whereas faster speech was associated with steeper decrement for the OC and PD groups and improvement for the DBS group. This interaction can be seen in Figure 26, and is captured by significant three-way interactions between group, rate, and vowel height, reported in Table 18 in Appendix E.

Specifically, this pattern is captured by the three-way interaction was significant for the OC vs. PD and DBS contrast in slow speech (S2: $\hat{\beta}$ = -0.718, p = 0.013; S3: $\hat{\beta}$ = -0.615, p = 0.01), as well as in faster speech (F2: $\hat{\beta}$ = -0.564, p = 0.023; the interaction was not significant for the fastest speech, F3: $\hat{\beta}$ = -0.464, p = 0.092). The three-way interaction was also significant for the PD vs. DBS contrast in faster speech (F2: $\hat{\beta}$ = -0.833, p = 0.005; the interaction was not significant for the fastest speech, F3: $\hat{\beta}$ = -0.373, p = 0.276).

4.1.3.2.4 Pairwise comparisons

The pairwise comparisons for the three-way interaction between group, rate, and vowel height in Figure 27 (and Table 18 in Appendix E) demonstrate that there were relatively few significant individual comparisons. For the OC group HNR is significantly better within high vowels in habitual speech versus both faster rates (F2, F3), and the slowest rate is significantly better than the fastest rate (S3 vs. F3). For the PD group, only one comparison was significant: habitual speech was associated with better HNR than the fastest speech (H1 vs. F3). For both high and low vowels produced by the DBS group, HNR is significantly worse in the slowest speech compared to the faster speech (S3 vs. F2), and in the slowest speech compared to habitual speech (H1 vs. S3).

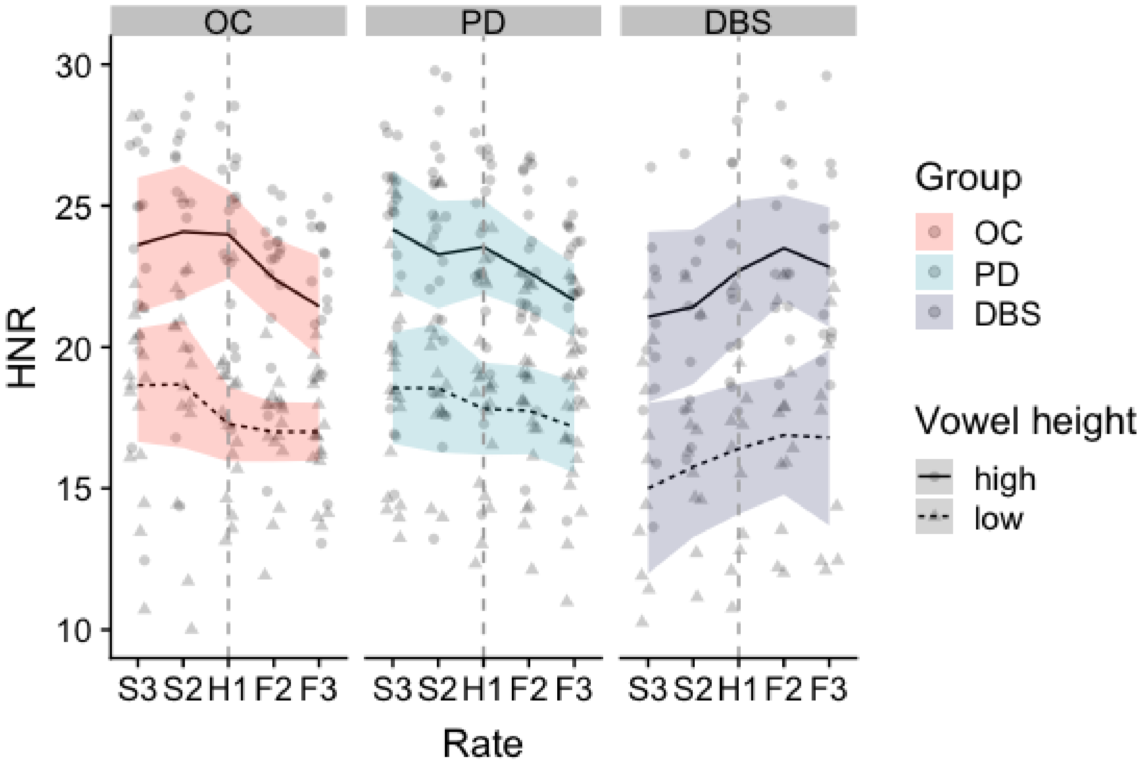


Figure 26: Harmonics-to-noise ratio by speaker group, speech rate, and vowel height, averaged across participants. Shaded band represents the 95% confidence interval. Points represent participant averages.

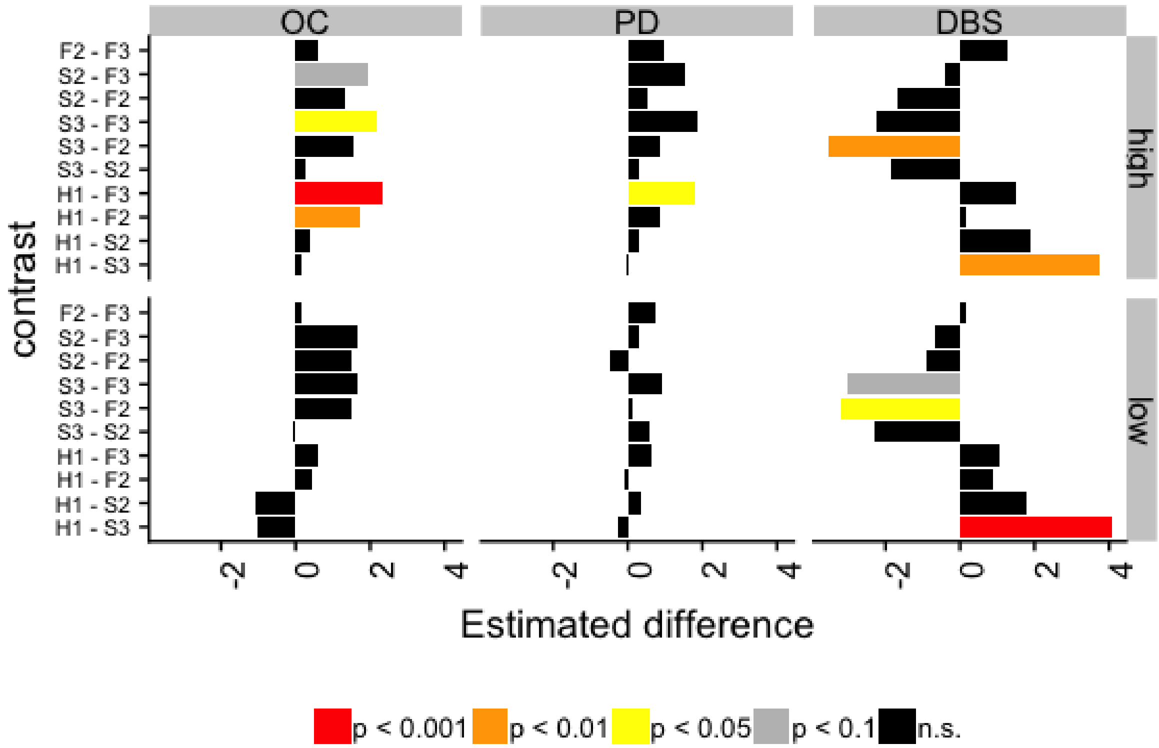


Figure 27: Harmonics-to-noise ratio pairwise comparisons for the three-way interaction between group, speech rate, and vowel height. X-axis represents estimated difference of the mean for the model. P-values were adjusted using the Tukey HSD method.

Rate by sex interactions

Overall, females demonstrated better voice quality at their habitual rates of speech compared to all modified rates, while males demonstrated better HNR in slow speech and worse HNR in fast speech. This is reported in Figure 28. This pattern was significant for all rate comparisons (S3 vs. H1: $\hat{\beta}$ = -1.677, p <0.001; S2 vs. H1: $\hat{\beta}$ = -0.925, p = 0.007; F2 vs. H1: $\hat{\beta}$ = -0.554, p = 0.05; F3 vs. H1: $\hat{\beta}$ = -0.891, p = 0.016).

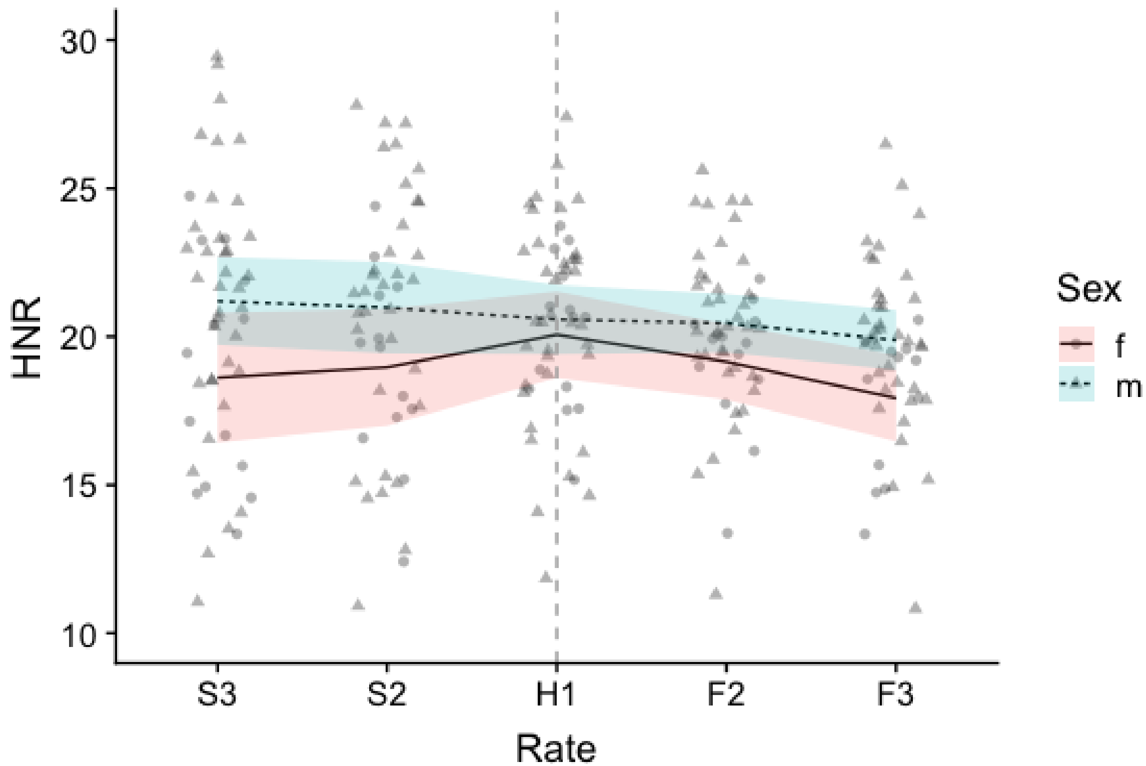


Figure 28: Harmonics-to-noise ratio by speech rate and speaker sex, averaged across participants. Shaded band represents the 95% confidence interval. Points represent participant averages.

Group, rate, sex interactions

While including the three-way interaction between group, rate, and sex improved the model fit, this interaction is not discussed here for simplicity. No pairwise comparisons between group, rate, and sex are reported because of the unbalanced sex distribution within and across groups.

4.2 Experiment 2: Transcription of nonce words

In this section, inter- and intra-reliability for the transcription task is reported first, followed by model results for the stop consonant accuracy and vowel accuracy analyses.

4.2.1 Inter-rater and intra-rater reliability

To reiterate, eight listeners came in over the course of four to five individual listening sessions and transcribed recordings of the sentences “Please say ___ again” presented in noise at +3dB SNR. Listeners were asked to transcribe the word in the blank and were given strict spelling criteria. They were told that words would always be of the form “aCVd”, where *C* was any possible consonant and *V* was any possible monophthong vowel of English. 20% of the files were repeated at random in order to calculate reliability. Each listener heard approximately 8500 sentences in total over the five sessions.

As stated in the Methods (Section 3.3.1.6), inter- and intra-rater reliability were calculated using Cohen’s kappa (Cohen, 1960). For ease of interpretation, both point-by-point agreement and Cohen’s kappa values are reported in Table 7. Inter-rater reliability was computed on a random sampling of 100 utterances for each listener pair (this was done because not all listeners heard the same overlap of utterances).

Average inter-rater word-level kappa was 0.358 (range: 0.02 to 0.56; agreement: 71.417); consonant-level kappa was 0.408 (range: 0.22 to 0.6; agreement: 74.292); and vowel-level kappa was 0.365 (range: 0.12 to 0.75; agreement: 74.375), indicating *fair* to *moderate* average agreement ranging from *slight* to *substantial* for individual listener pairs (Landis & Koch, 1977).

Average intra-rater word-level kappa was 0.556 (range: 0.508 to 0.625; agreement: 77.971); consonant-level kappa was 0.61 (range: 0.555 to 0.661; agreement: 82.059), and vowel-level kappa was 0.515 (range: 0.441 to 0.58; agreement: 81.394), indicating a *moderate* to *substantial* average and range (Landis & Koch, 1977).

Table 7: Intra-rater reliability for each listener: Nonsense word transcription task. Point-by-point agreement and Cohen’s kappa are listed for 1) whole word accuracy (i.e., consonant and vowel), 2) consonant accuracy, and 3) vowel accuracy

Listener	Word		Consonant		Vowel	
	Agreement	Kappa	Agreement	Kappa	Agreement	Kappa
1	78.790	0.566	83.330	0.606	84.610	0.561
2	81.350	0.625	83.970	0.661	81.630	0.580

Listener	Word		Consonant		Vowel	
	Agreement	Kappa	Agreement	Kappa	Agreement	Kappa
3	78.150	0.563	82.550	0.611	81.380	0.521
4	76.840	0.535	82.790	0.614	78.820	0.441
5	75.360	0.508	78.820	0.555	83.570	0.461
6	77.280	0.532	79.080	0.568	77.070	0.513
7	76.500	0.531	82.360	0.604	82.570	0.551
8	79.500	0.590	83.570	0.661	81.500	0.494

4.2.2 Transcription accuracy

The analysis for this section was run on the same subset of the data reported in 4.1: items from the nonce word task containing stop consonants (items 1 - 24). This amounted to 24 items per participant per rate condition (n = 7) per participant (n = 68 across all four groups), or 11,424 trials in total. Most but not all speakers were able to complete all trials, as noted in the methods. Utterances that were somehow unpresentable in some way (e.g., recording cut off, participant misspoke) were discarded with 11,241 unique utterances included in the final analysis (n excluded = 183; 1.6%). The responses for the first two listeners for each speaker were retained, resulting in 22,482 observations in the data set analyzed here.

4.2.3 Stop consonant accuracy

4.2.3.1 Final model

Transcription accuracy of stop consonants was modelled as a function of *group*, *rate*, *voicing*, and *place of articulation* using logistic mixed effects regression. All fixed effects and two-way interactions were entered. Three-way interactions involving both group and rate were also entered. A fixed effect of *listener* was added as a covariate to account for differences across transcribers.

As with the acoustics models in Section 4.1, the base model included fixed effects of *group* and *rate* as well as their interaction. *Listener* was included as a covariate in the base model as well. Random effects included by-participant random intercepts and slopes for each rate contrast, as well as by-item random intercepts and slopes for each group contrast.

The final model included fixed effects of group, rate, voicing, and PoA. Two-way interactions that improved the model fit included group x voicing and group x PoA. Two-way interactions between rate, voicing, and PoA did not improve the model fit. The three-way interaction between group x rate x voicing improved the model, but group x rate x PoA did not. The fixed effects structure can be summarized as:

Stop consonant accuracy ~ group + rate + voicing + PoA +
group x rate + group x voicing + group x PoA +
group x rate x voicing + listener

4.2.3.2 Main effects

Model results reporting follows that of Experiment 1. Main effects of fixed effects terms are reported here, and interactions are reported below. Coefficients are reported in Table 19 in Appendix F. Post-hoc pairwise comparisons for the two-way interaction between group and rate are reported in Table 20 in Appendix F and in Figure 29. Figure 31 displays the empirical data for patterns of stop consonant transcription accuracy by speaker group, speech rate, and stop consonant voicing.

The model coefficients may be interpreted in the following way: a *positive* estimate in the reported model coefficients indicates that the reference level for that contrast was associated with greater transcription accuracy.

Speaker group

When all other variables were held at their average values, the healthy control groups were transcribed with higher accuracy than the clinical groups, as indicated by significant positive effects for these contrasts (YC vs. OC, PD, DBS: $\hat{\beta}$ = 1.671, p <0.001; OC vs. PD, DBS: $\hat{\beta}$ = 0.984, p <0.001), but the PD and DBS groups did not significantly differ from one another ($\hat{\beta}$ = 0.089, p = 0.735).

Speech rate

In the main model, each modified speech rate was compared with the habitual rate when all other variables are held constant (consistent with what was done in Section 4.1).

Overall, the two slower rates (S3, S2) were associated with *worse* (lower) transcription accuracy compared to the habitual rates (S3vsH1: $\hat{\beta}$ = -0.282, p = 0.009; S2vsH1: $\hat{\beta}$ = -0.222, p = 0.019), while the faster rates (F2, F3) did not significantly differ from habitual (F2vsH1: $\hat{\beta}$ = -0.046, p = 0.564; F3vsH1: $\hat{\beta}$ = -0.065, p = 0.608).

Voicing

Voiceless stops were transcribed with greater accuracy than voiced stops ($\hat{\beta}$ = -0.202, p = 0.002).

Place of articulation

Bilabial stops (/b, p/) did not differ from non-bilabial stops (/t, d, k, g/) ($\hat{\beta}$ = -0.007, p = 0.944), but alveolar stops (/t, d/) were transcribed with greater accuracy than velar stops ($\hat{\beta}$ = 0.468, p <0.001).

4.2.3.3 Interactions and pairwise comparisons

Group by rate interactions

The majority of the group x rate comparisons were not significant, indicating that speech rate adjustments impacted consonant intelligibility across the groups in similar ways when all other factors were equal. An exception to this was that the YC group demonstrated greater decrement in consonant intelligibility in the fastest speech (F3) compared to habitual, demonstrated by a significant interaction for the YC vs. OC, PD, DBS by F3 ($\hat{\beta}$ = -0.564, p = 0.029). Post-hoc pairwise comparisons are reported in Figure 29, which emphasize this point: most of the contrasts between rates were not significant.

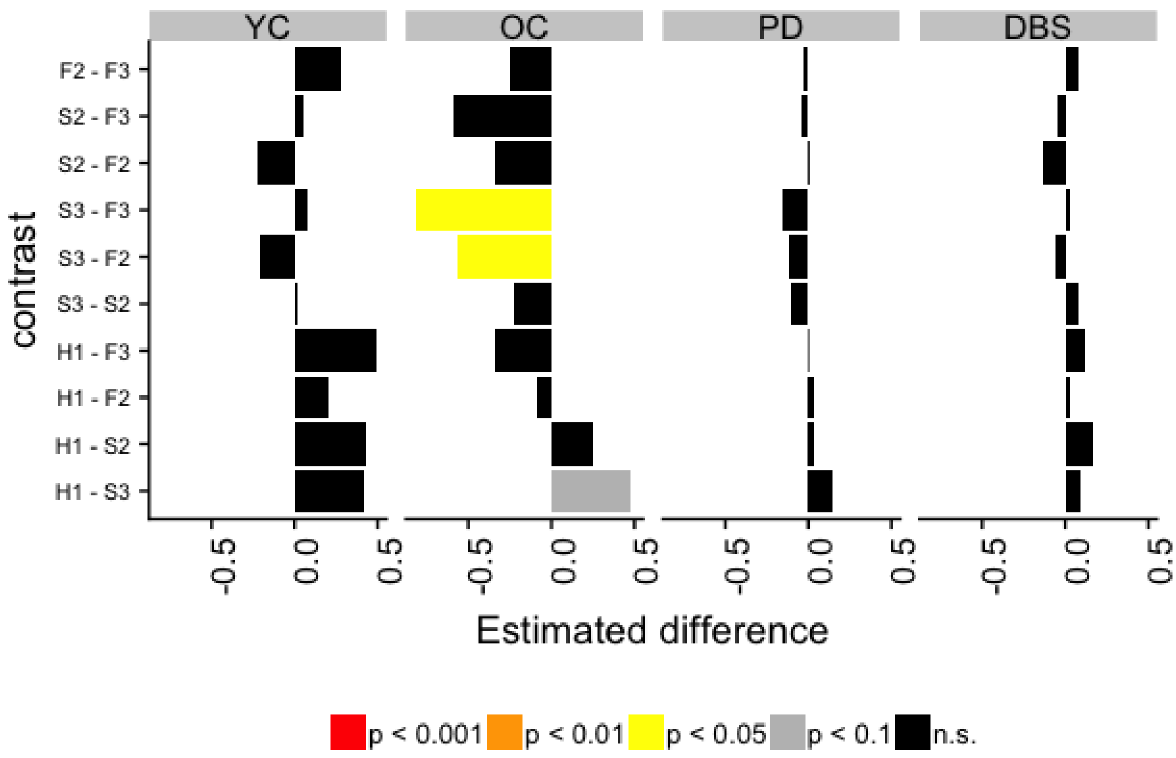


Figure 29: Stop consonant accuracy pairwise comparisons for the two-way interaction between group and speech rate. X-axis represents estimated difference of the mean for the model. Results are given on the log-odds ratio scale and are averaged over the levels of voicing, PoA, and listener. P-values were adjusted using the Tukey HSD method.

Group by voicing interactions

While all groups were transcribed with greater accuracy when producing voiceless rather than voiced stops, this difference was greater for the OC, PD, and DBS groups compared to the YC group. This is captured by a marginally significant interaction between the YC vs. OC, PD, DBS by voicing contrast ($\hat{\beta}$ = 0.211, p = 0.05). There was also a trend for the difference between voiced and voiceless stops to be greater for the OC group compared to the clinical groups, as captured by a marginally significant interaction for the OC vs. PD, DBS by voicing interaction ($\hat{\beta}$ = -0.16, p = 0.072). The PD and DBS groups did not differ from one another ($\hat{\beta}$ = -0.158, p = 0.121).

Group by PoA interactions

All possible two-way interactions between group and PoA were significant. Empirical patterns can be seen in Figure 30. The YC group was transcribed with greatest accuracy for alveolar stops compared to bilabial and velar stops, and this difference was greater than for the other three groups. This is captured by negative interactions with the bilabial vs. non-labial contrast for the YC vs. OC, PD, DBS comparison (i.e., bilabials for the YC group were transcribed with lower accuracy than alveolar and velars combined; YC vs. OC, PD, DBS: $\hat{\beta} = -0.502$, $p < 0.001$), and a *positive* interaction for the YC vs. Rest x Alveolar vs. Velar comparison ($\hat{\beta} = 0.445$, $p < 0.001$). The OC, PD, and DBS groups differed from the YC group, but were similar to each other. The *degree* to which stop intelligibility differed across PoAs for each group differed, however (e.g., flatter slopes of change for the clinical groups) This is captured by significant positive interactions for both PoA contrasts for the OC vs. PD and DBS comparison (bilabial vs. non-labial: $\hat{\beta} = 0.189$, $p = 0.03$; alveolar vs. velar: $\hat{\beta} = 0.293$, $p = 0.004$) and the PD vs. DBS comparison (bilabial vs. non-labial: $\hat{\beta} = 0.297$, $p = 0.004$; alveolar vs. velar: $\hat{\beta} = 0.247$, $p = 0.041$) indicate that the overall pattern across PoAs for these three groups was similar, but the degree to which intelligibility differed across PoAs for each group differed (e.g., flatter slopes of change for the clinical groups).

To summarize, bilabial stops were transcribed with significantly lower accuracy than alveolar and velar stops combined for the YC group, but not for the other groups. Conversely, all groups demonstrated lower intelligibility for velar stops overall, but this difference was greatest for the YC group. Alveolar stops were transcribed with greatest accuracy for the YC group, whereas this distinction was not as clear for the OC, PD, and DBS groups. Velar stops were transcribed with the lowest accuracy for all groups. Figure 30 demonstrates these relationships.

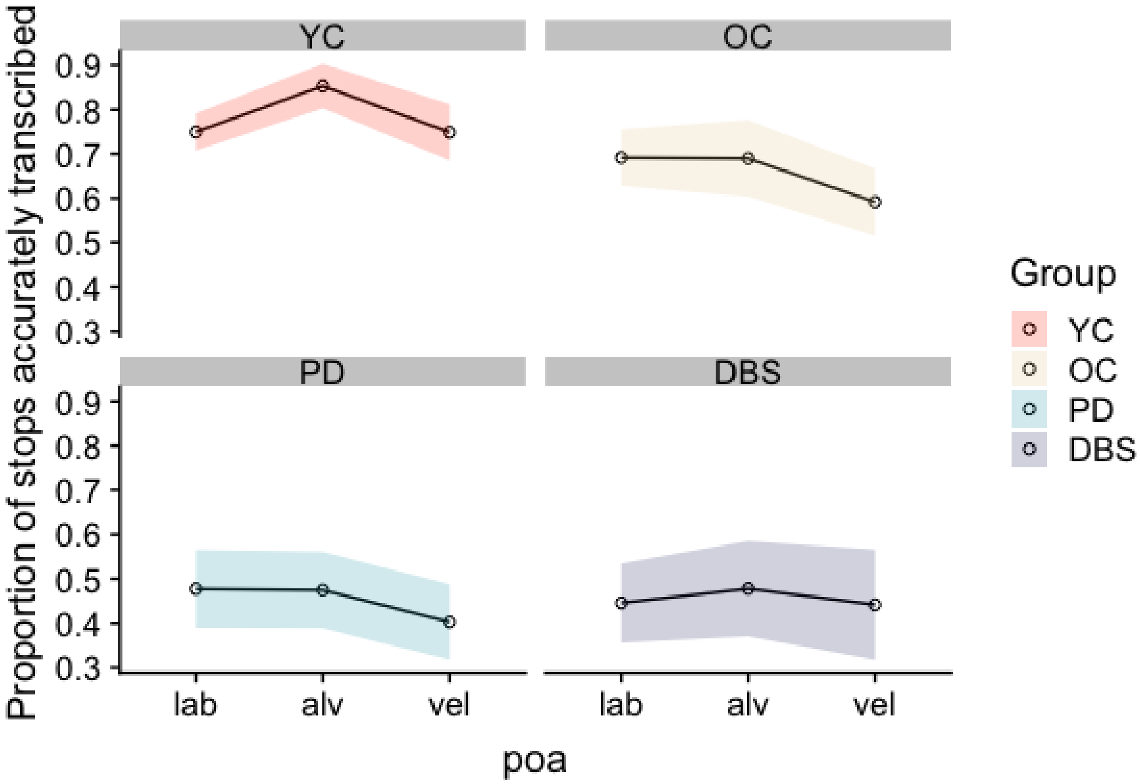


Figure 30: Proportion of correctly transcribed stop consonants by speaker group and consonant place of articulation. Averages first aggregated by listeners and participants. Shaded band represents the 95% confidence interval.

Group, rate, and voicing interaction

Though the inclusion of the three-way interaction between group, rate, and voicing improved the model fit, none of the main model comparisons were significant. Pairwise comparisons are thus only reported for the two-way interaction between group and rate (Figure 29).

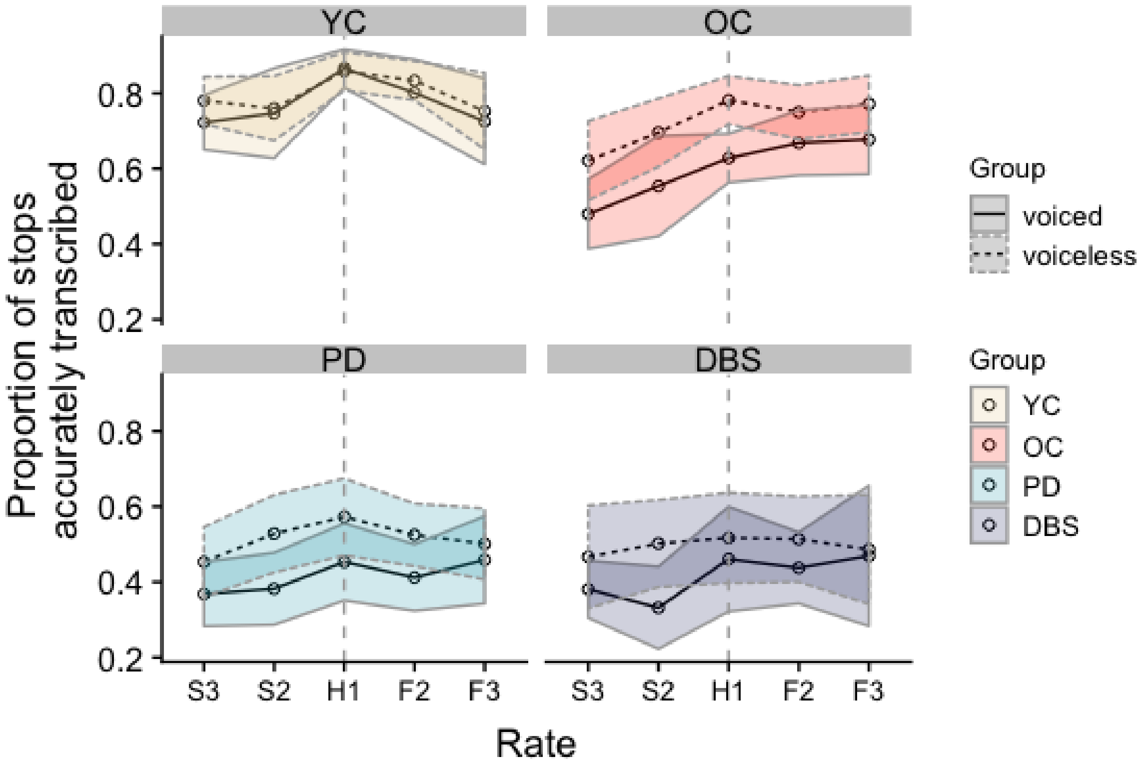


Figure 31: Proportion of correctly transcribed stop consonants by speaker group, speech rate, and consonant voicing. Proportions averaged by listeners and participants. Shaded band represents the 95% confidence interval.

In summary, consonant intelligibility for the four speaker groups differed in a predictable manner (YC > OC > PD/DBS), but rate of speech did not differentially impact how well each of the groups was understood (in terms of stop consonant identification). Notably, slow speech was significantly associated with *worse* intelligibility and the highest proportions of stop accuracy were in habitual speech. Voiceless, alveolar stops (/t/) were transcribed with the highest accuracy, and flatter rates of change were observed for the PD and DBS groups compared to the controls.

4.2.4 Vowel accuracy

4.2.4.1 Final model

Vowel accuracy was modelled as a function of *group*, *rate*, and *vowel backness*. *Listener* was also included as a covariate, as was done for the consonant transcription analysis. The iterative model building procedure was the same as described previously. All terms and interactions entered into the model were associated with significant model improvement and so none were dropped. The final model thus included fixed effects of *group*, *rate*, *vowel backness*, and *listener*, all two-way interactions between *group*, *rate*, and *vowel backness* as well as the three-way interaction between *group*, *rate*, and *vowel backness*. The random effects structure was identical to that described in the stop accuracy models.

The fixed effects structure for the final model can be summarized as:

> Vowel accuracy ~ group + rate + vowel backness + group x rate + group x vowel backness + rate x vowel backness + group x rate x vowel backness + listener

4.2.4.2 Main effects

Figure 32 reports the empirical data for vowel accuracy. Model coefficients are in Table 21 in Appendix F. Pairwise comparisons for the three-way interaction between *group*, *rate*, and *vowel backness* are in Figure 33 as well as in Table 22 in Appendix F.

Speaker group

The YC group was transcribed with greater accuracy than the other three groups ($\hat{\beta}$ = 0.905, p <0.001), and the OC group was transcribed with greater accuracy than the PD and DBS groups ($\hat{\beta}$ = 1.011, p <0.001). The PD and DBS groups did not differ from each other ($\hat{\beta}$ = 0.337, p = 0.175). This can be summarized as: YC > OC > PD/DBS.

Speech rate

The slower rates were both significantly associated with poorer vowel accuracy compared to talkers' habitual rates of speech (S3 vs. H1: $\hat{\beta}$ = -0.268, p = 0.004; S2 vs. H1: $\hat{\beta}$ = -0.253, p = 0.008). Vowel transcription accuracy in *fast* speech did not significantly differ from habitual speech (F2 vs H1: $\hat{\beta}$ = -0.047, p = 0.62; F3 vs. H1: $\hat{\beta}$ = -0.216, p = 0.105).

Vowel backness

Front vowels were transcribed with poorer accuracy than back vowels ($\hat{\beta}$ = -0.607, p <0.001).

4.2.4.3 Interactions

Empirical data are reported in Figure 32, and pairwise comparisons for the three-way interaction between group, rate, and vowel backness are reported in Figure 33 and in Table 22 in Appendix F.

Group by rate

The YC group demonstrated a greater difference in the slow versus habitual rates compared to the OC, PD, and DBS groups combined, as evidenced by significant positive interactions for the YC vs. Rest comparison at the slower (S2: $\hat{\beta}$ = 0.471, p = 0.04) and slowest (S3: $\hat{\beta}$ = 0.507, p = 0.014) rates. This was mainly driven by the finding that the OC group, conversely, demonstrated *lower* intelligibility in the slowest rate compared to the clinical groups, indicated by a significant *negative* interaction for the OC vs. PD, DBS comparison at the S3 rate ($\hat{\beta}$ = -0.536, p = 0.01). With other variables held constant, the YC essentially were *as intelligible* in slow speech, the OC group was much *less* intelligible in slow speech, and the PD and DBS groups were *somewhat* less intelligible in slow speech compared to the controls and compared to their habitual rates. The PD group had *worse* intelligibility in both fast *and* slow speech rates, whereas intelligibility was overall unchanged for the DBS group. The difference between the PD and DBS groups only reached significance for the faster speech (F2: $\hat{\beta}$ = -0.583, p = 0.011).

Group, rate, and vowel backness

Unlike for the stop transcription, vowel transcription accuracy demonstrated significant differences for all groups across speech rates, and these patterns differed for front and back vowels. Patterns described here are visible in Figure 32 and are followed by the results of the pairwise comparisons in order to describe changes between each rate for each group.

In slower speech, compared to habitual, both front and back vowels were transcribed with similar accuracy for the PD groups. The OC group, on the other hand, showed the opposite trend (more similarity at habitual rates compared to slower speech). This is reflected by a significant negative interaction for the OC vs. PD, DBS comparison (S2: $\hat{\beta}$ = -0.4, p = 0.013;

S3: $\hat{\beta}$ = -0.329, p = 0.025), and a significant positive interaction for the PD vs. DBS (S3: $\hat{\beta}$ = 0.395, p = 0.006; S2 was marginally significant: $\hat{\beta}$ = 0.308, p = 0.052).

The PD group also demonstrated greater difference in their front and back vowel intelligibility rates in faster speech compared to the DBS group (F2: $\hat{\beta}$ = 0.049, p = 0.747), mainly driven by a steeper decline of front vowel intelligibility in fast speech. None of the comparisons between the YC and the other groups were significant for this interaction.

These patterns are visible in Figure 32 as the gap between front and back vowel accuracy in slow compared to habitual speech widens for controls, closes for PDs, and remains relatively stable for the DBS and YC groups.

The PD and DBS groups front vowel productions do not benefit from slow speech, whereas front vowel accuracy declines. The OC group shows similarly high accuracy for both vowel categories at habitual rates, but front vowel accuracy declines at a steeper rate than for back vowels.

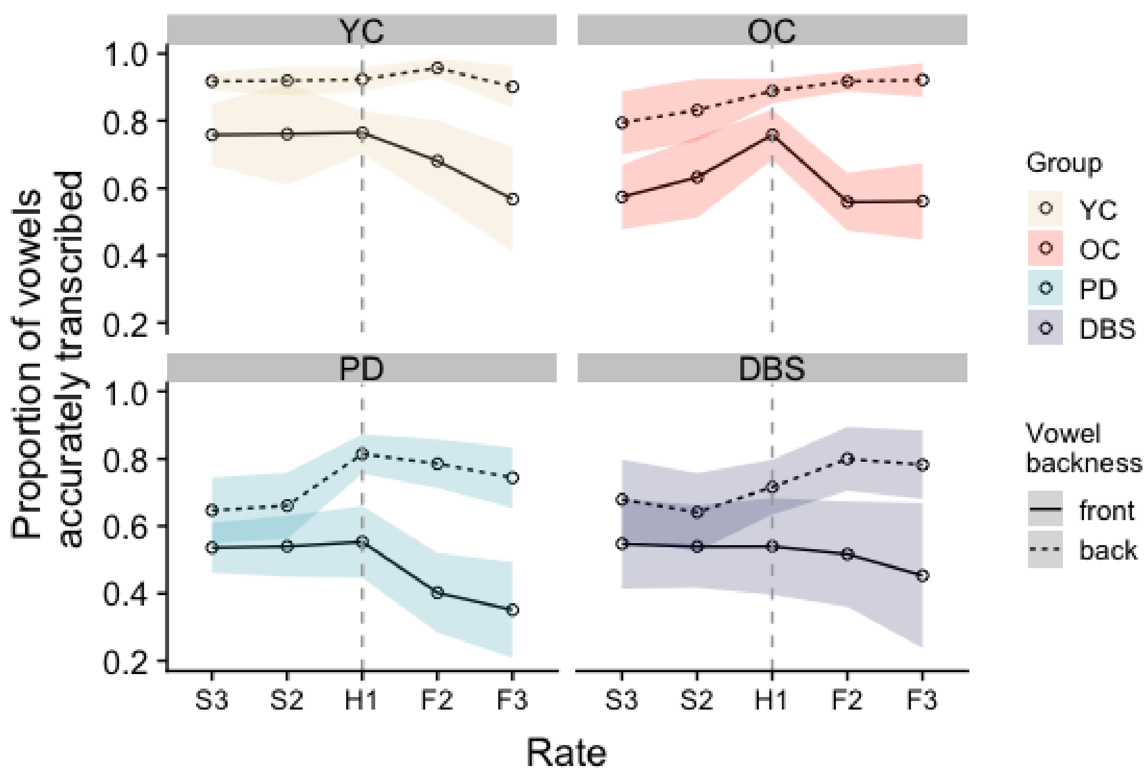


Figure 32: Proportion of correctly transcribed vowels by rate and vowel backness. Averages first aggregated by listeners and participants. Vertical dotted line represents habitual rate. Shaded band represents the 95% confidence interval.

4.2.4.4 Pairwise comparisons

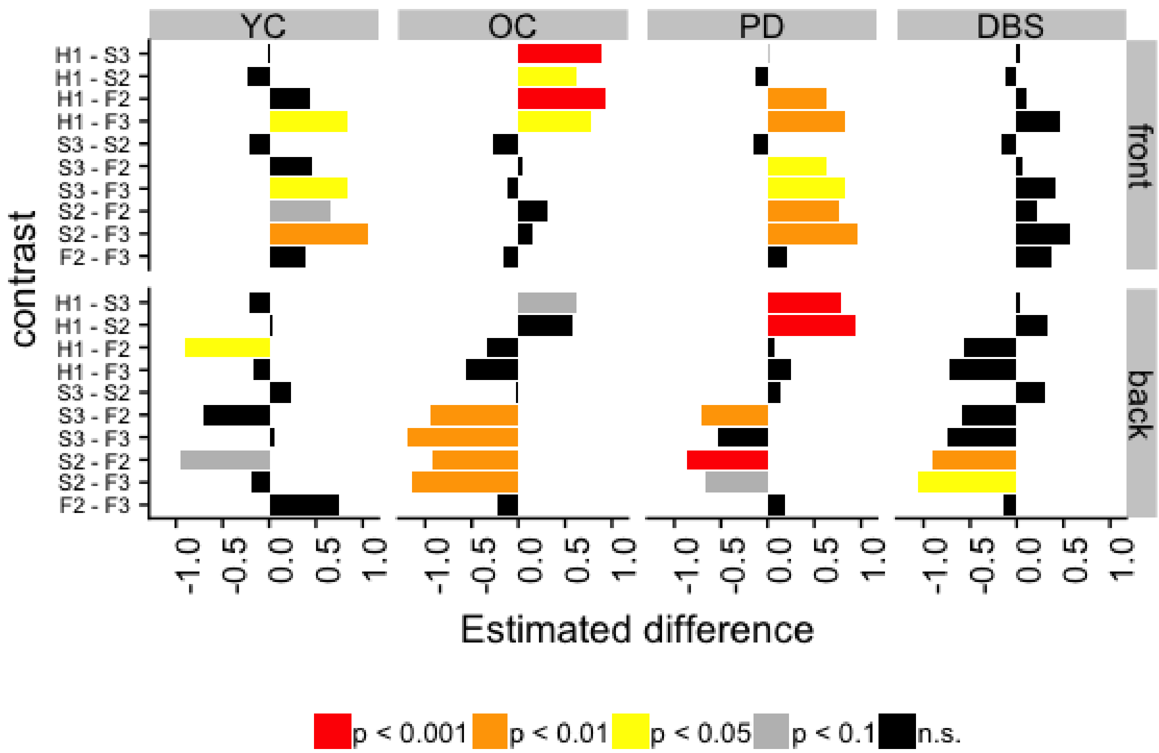


Figure 33: Vowel transcription accuracy pairwise comparisons for the three-way interaction between group, speech rate, and voicing, illustrating differences between rates. X-axis represents estimated difference of the mean for the model. Results are given on the log-odds ratio and averaged over the levels of listener. P-values were adjusted using the Tukey HSD method.

The pairwise comparisons for the three-way interaction are reported below in order to better understand specific differences between the different speech rates for each group and vowel category. These will be discussed in terms of *habitual versus modified*, *slow versus fast*, and *within slow or fast rates*. Significant pairwise comparisons are reported, and a report of the effect size, direction, and significance can be seen in the accompanying Figure 33. The empirical data reported in Figure 32 aids in these interpretations.

In the pairwise comparison figures, a positive estimate indicates that the contrast specified on the left was associated with *greater* accuracy than that on the right (e.g., a positive estimate for the H1-S3 contrast for the OC group in the top panel indicates that front vowels were transcribed with greater accuracy in habitual speech than the slowest speech for the OC group; the red color indicates that this was significant at $p < 0.001$).

Habitual rate versus modified rates

Compared to their habitual speech, vowel accuracy for the YC group did not significantly differ in any of the modified speech rates with two exceptions: front vowels were more accurately transcribed in habitual speech versus the fastest speech (F3), while back vowels were transcribed with greater accuracy in faster speech (F2) compared to habitual.

Front vowels for the OC group were transcribed with greatest accuracy in the habitual condition compared to both slower and faster rates, but back vowels did not differ.

Overall, the PD group was most intelligible (highest proportion of vowels accurately transcribed) in habitual speech, but the pattern of significance differed for front and back vowels. Specifically, front vowel accuracy demonstrated a steeper decline in fast speech (F2 and F3), while back vowel accuracy declined more in slow speech (S2 and S3).

There were no significant comparisons between habitual rates of speech and any of the modified rates of speech for the DBS group.

Slower versus faster speech

Within slow or fast speech, vowel accuracy did not demonstrate significant differences: slower and slowest (S2 versus S3) did not differ from one another for any group or either vowel position, nor did the faster and fastest rates (F2 versus F3).

A comparison across slow and fast speech demonstrated the following patterns. In general, front vowels were transcribed more accurately in slow versus fast speech, whereas back vowels were transcribed more accurately in fast versus slow speech. Significant comparisons are reported below.

For *front* vowels, the above-mentioned pattern reached significance only for the YC and PD groups. Specifically, both slower and slowest speech rates were more accurately transcribed than the *fastest* speech rate for both groups (S3 > F3 and S2 > F3). Both slow rates were also associated with greater accuracy for front vowels compared to *faster* speech (S3 > F2 and S2 > F2), though this comparison only reached significance for the PD group.

For *back* vowels, the pattern was reversed; namely, slower speech was, in general, associated with **poorer** transcription accuracy compared to faster speech. These comparisons were significant for all groups except for the YC group, who did not demonstrate significant changes in faster speech for back vowels. The slowest speech (S3) differed from faster speech (F2) for the OC and PD groups, and from fastest speech (F3) for the OC group only. There were no differences between the slowest speech and any of the fast rates for the DBS group. The slower speech rate (S2) was associated with significantly poorer accuracy than faster speech (F2) for the OC, PD, and DBS groups, as well as than fastest speech (F3) for the OC and DBS groups.

In summary, the modified speech rates did not improve vowel transcription accuracy compared to habitual speech. Across the modified speech rates, slower speech was generally associated with better identification of front vowels (/i, æ/) and poorer identification of back vowels (/u, ɑ/). For the most part, this was consistent across group, and in particular for the PD group. An interesting finding is that the DBS group demonstrated better intelligibility in fast but not slow speech, and only for back vowels.

4.3 Experiment 3: Speech intelligibility estimation

4.3.1 Inter-rater and intra-rater reliability

As described in the methods (Section 3.3.2.5), listener reliability for the speech intelligibility estimation tasks were computed using the intraclass correlation coefficient (ICC; Koo & Li, 2016). Separate coefficients were computed for each task. Inter-rater reliability was calculated as the average consistency in a two-way random model (ICC 2, k), and intra-rater reliability was calculated using average agreement in a two-way mixed model (ICC 3, k).

Average inter-rater reliability for the sentence task was 0.889 (95% CI: 0.878 - 0.898), and 0.938 (95% CI: 0.921 - 0.951) for the conversation task. This can be interpreted as *good* and *excellent* inter-rater reliability, respectively (Koo & Li, 2016).

Average intra-rater reliability for the sentence task was *good* (mean: 0.872, range: 0.824 - 0.933) and *excellent* for the conversation task (mean: 0.934, range: 0.894 - 0.983) (Koo & Li, 2016). Intra-rater reliability scores for each listener and task are presented in Table 8 and Table 9[21].

Table 8: Intra-rater reliability for each listener: Sentence rating task

Listener	ICC	F	df1	df2	p	lower bound	upper bound
1	0.831	5.929	540.000	540.000	<0.001	0.800	0.858
2	0.899	9.916	543.000	543.000	<0.001	0.881	0.915
3	0.933	14.955	273.000	273.000	<0.001	0.915	0.947
4	0.824	5.684	273.000	273.000	<0.001	0.777	0.861
5	0.884	8.653	273.000	273.000	<0.001	0.853	0.909
6	0.861	7.188	273.000	273.000	<0.001	0.824	0.890

Note: ICC = Intraclass correlation coefficient (ICC 3,k).

Table 9: Intra-rater reliability for each listener: Conversation rating task

Listener	ICC	F	df1	df2	p	lower bound	upper bound
1	0.983	57.921	47.000	47.000	<0.001	0.969	0.990
2	0.945	18.061	47.000	47.000	<0.001	0.901	0.969
3	0.895	9.546	47.000	47.000	<0.001	0.813	0.941
4	0.911	11.215	47.000	47.000	<0.001	0.841	0.950
5	0.894	9.410	47.000	47.000	<0.001	0.810	0.940
6	0.975	40.240	47.000	47.000	<0.001	0.956	0.986

[21] The higher degrees of freedom for Listeners 1 and 2 in Table 8 reflect that they were presented with 20% instead of 10% of files for reliability purposes. The amount repeated was later reduced to minimize the task time.

Listener	ICC	F	df1	df2	p	lower bound	upper bound

Note: ICC = Intraclass correlation coefficient (ICC 3,k).

4.3.2 Intelligibility estimation results

To reiterate, the purpose of this analysis was to model listener ratings of intelligibility as a function of speaker group, speech rate, and task complexity. Figures 34 - 36 show the empirical data across these variables.

Two analyses were done. The first modelled *sentence reading intelligibility* as a function of group, rate, and sentence length (i.e., number of words). The second modelled *conversational intelligibility* as a function of group and rate. This was run as a separate model because speech rate could not be measured in the same way for sentence reading and conversation. As such, the rate metric entered in the conversational intelligibility model was the rate *condition*, rather than the proportional rate. Recall that the conversational task was always elicited as the last item in each rate condition, and therefore participants were maximally habituated to it. Conversational samples for this analysis (approximately 10 to 20 seconds per participant and condition) were also specifically selected to be maximally representative of the experimental rate condition. Therefore, while this section differs from the others in that rate was not treated as a true proportional rate, including rate condition as the rate metric is a valid approach.

Additionally, because the sentence and conversational samples could not be included in the same model, a Wilcoxon-signed rank test was performed on the intelligibility ratings for each task. There was no significant difference between the two tasks (z = -0.259, p = 0.795)[22].

Both models included random by-participant intercepts with random slopes for the rate contrasts (categorical proportional rate for the sentences, as used previously, and rate *condition* for conversation).

[22] Despite the finding that the two tasks were overall similar, separate mixed model analyses were still conducted in order to explore effects of group and rate in a more robust manner.

The main results of each of these predictor variables is first reported in this section. Coefficients for all main effects and interactions are reported in Appendix G.

Note that a negative estimate in the reported model coefficients indicates that the contrast label on the left was associated with *higher* intelligibility ratings than the label on the right[23].

4.3.2.1 Sentence reading intelligibility

4.3.2.1.1 Final model

The final model for sentence reading included main fixed effects of *group, speech rate,* and *sentence length*. Two-way interactions between rate and sentence length were included. The two-way interaction between group and sentence length did not improve the model fit, nor did the three-way interaction. These terms were thus excluded from the final model. Sentence length was treated as a categorical variable. The fixed effects structure for the final model can be summarized as:

log(100 - Intelligibility) ~ group + rate + sentence length +
group x rate + rate x sentence length

4.3.2.1.2 Main effects

Empirical data for the sentence intelligibility task appear in Figures 34 through 35. Model coefficients are reported in Table 23 in Appendix G. The two-way interaction between group and rate was significant but no post-hoc pairwise comparisons are reported.

Speaker group

All three group contrasts were significant and may be interpreted in the following way: the YC group was rated as most intelligible (YC vs. OC, PD, DBS: $\hat{\beta}$ = -0.798, p <0.001), followed by the OC group (OC vs. PD, DBS: $\hat{\beta}$ = -1.092, p <0.001), with the DBS group being rated as least intelligible (PD vs. DBS: $\hat{\beta}$ = -0.469, p = 0.014).

---

[23] This negative relationship reflects the fact that the response variable, intelligibility, was subtracted from a constant before being log-transformed. This pattern differs from previous results sections in this thesis and so is reported here for clarity.

Speech rate

All rate comparisons were also significant. Overall, both slower rates (S3, S2) were rated as *more* intelligible than habitual (S3: $\hat{\beta}$ = -0.303, p <0.001; S2: $\hat{\beta}$ = -0.391, p <0.001), and both faster rates (F2, F3) were rated as significantly *less* intelligible than habitual (F2: $\hat{\beta}$ = 0.446, p <0.001; F3: $\hat{\beta}$ = 0.879, p <0.001). In other words, with the other variables held at their average values, the slower the speech, the more intelligible it was rated.

Sentence length

In most cases, longer sentences were associated with lower intelligibility, as captured by significant negative contrasts for each of the sentence length comparisons (5-words vs. 6-words: $\hat{\beta}$ = -0.175, p <0.001; 5 to 6 words vs. 7-words: $\hat{\beta}$ = -0.179, p <0.001; 5 to 8 words vs. 9-words: $\hat{\beta}$ = -0.035, p = 0.003; 5 to 9 words vs. 10-words: $\hat{\beta}$ = -0.076, p <0.001). An exception to this pattern was the comparison between the shorter sentences and the 8-word sentence (5 to 7 words vs. 8-words: $\hat{\beta}$ = 0.066, p <0.001). Here the significant negative contrast indicates that the 8-word sentence was rated as more intelligible overall compared to the shorter (5 to 7-word) sentences.

Figure 34 A demonstrates that this pattern, pictured as a spike in intelligibility ratings for the 8-word sentences, was especially apparent for the PD and DBS groups. The decrease in intelligibility as sentence length increases is most obvious for the DBS group. Note however that this plot does *not* demonstrate a *significant* interaction between group and sentence length. This interaction did not demonstrate a significant improvement in the model fit during the iterative model building process and was thus excluded from the final model. Nevertheless, it is pictured here to demonstrate the variability across groups.

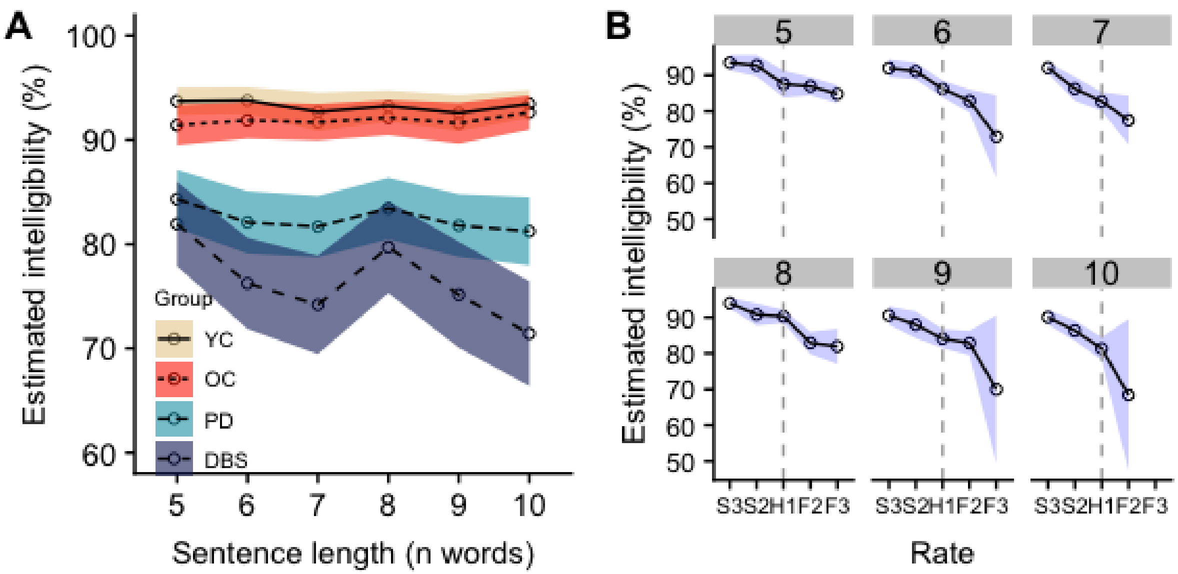


Figure 34: Sentence intelligibility by: A) sentence length and speaker group and B) speech rate and sentence length (individual panels correspond to distinct sentence lengths). All points were first aggregated by participants. In B, vertical line at H1 indicates habitual rate; individual panels represent each of the six sentence lengths (five to ten words). Shaded band represents the 95% confidence intervals.

4.3.2.1.3 Interactions

Group by rate interactions

The empirical data reporting this interaction appear in Figure 35. Model coefficients are reported in Table 23 in Appendix G. The pattern in sentence intelligibility across speech rates (higher in slow and poorer in fast speech) was roughly similar for the YC, OC, and PD groups, but to different degrees. Significant interactions at each rate contrast (S3, S2, F2, F3) for all YC vs. Rest comparisons mainly indicate that the effect was *stronger* for the YC group, presumably due to much less variability due to a tight cluster of high intelligibility ratings (S3: $\hat{\beta}$ = -0.296, p = 0.009; S2: $\hat{\beta}$ = -0.477, p <0.001); F2: $\hat{\beta}$ = 0.265, p = 0.029; F3: $\hat{\beta}$ = 0.541, p = 0.002).

For the OC vs. PD, DBS contrast, there were no significant interactions in the slow speech comparisons (S2: $\hat{\beta}$ = -0.101, p = 0.393; S3: $\hat{\beta}$ = -0.019, p = 0.877), indicating similar gains in intelligibility in slow speech from habitual speech for the OC group compared to the pooled PD and DBS groups. The OC group demonstrated a *greater decline* in intelligibility in fast speech, indicated by significant interactions for the OC vs. PD, DBS contrast for both F2 ($\hat{\beta}$ = 0.335, p = 0.012) and F3 ($\hat{\beta}$ = 0.491, p = 0.008). This should be considered with caution, however, given that the PD and DBS groups demonstrated differences in fast speech. That is, the OC group demonstrated greater decline compared to the pooled intelligibility of the PD and DBS groups, possibly because the opposite trends of the PD and DBS groups cancelled one another out.

Similarly, the PD and DBS groups did not differ in the slow speech comparisons (S2: $\hat{\beta}$ = -0.096, p = 0.488; S3: $\hat{\beta}$ = -0.21, p = 0.157), nor did they differ in faster speech (F2: $\hat{\beta}$ = 0.267, p = 0.095). At their *fastest* speech rate, however, intelligibility ratings for the PD group continued to decline, but actually were shown to *increase* for the DBS group. This difference was significant ($\hat{\beta}$ = 0.555, p = 0.01), though it should be noted that this interaction merely indicates that intelligibility for the PD group was *lower* in fastest compared to habitual and does not capture the change in direction. That can be seen from the empirical data in Figure 35, which shows that there was a trend for the DBS group to improve in both slow and fast speech compared to their habitual rate. Post-hoc pairwise comparisons were carried out to investigate this observed pattern; none of the pairwise comparisons for the DBS group were significant.

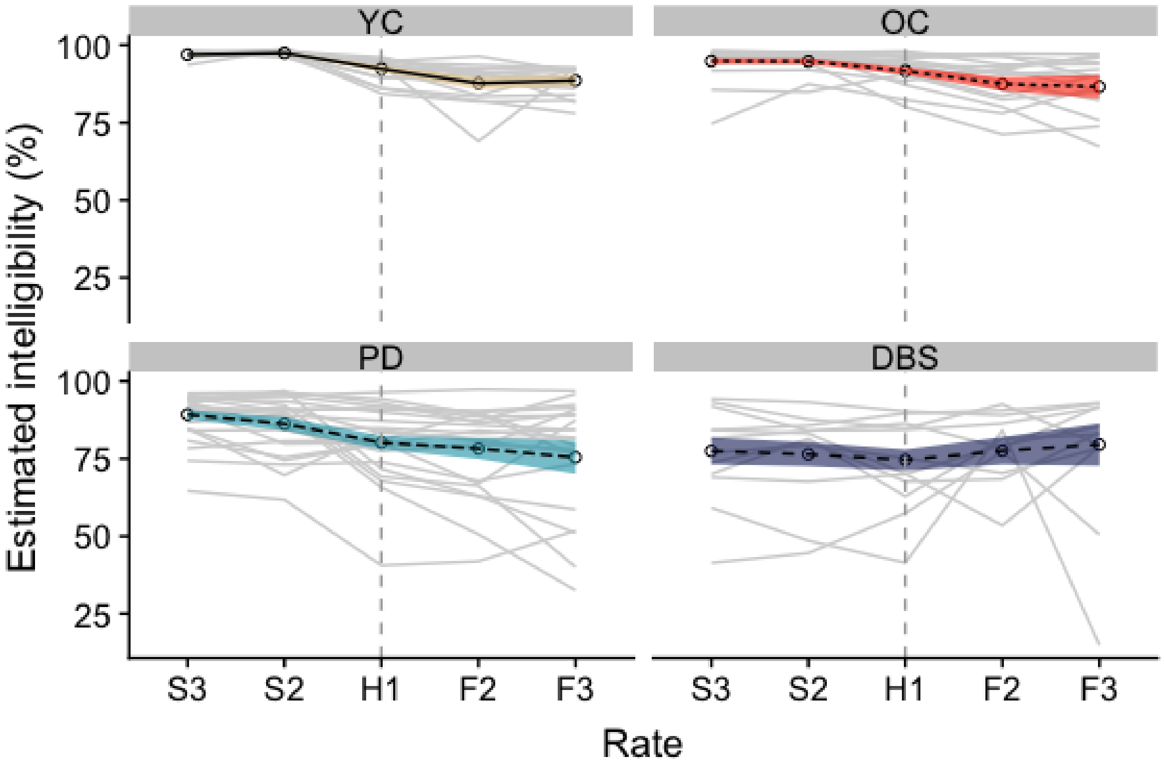


Figure 35: Sentence intelligibility by speaker group and rate. Vertical line at H1 represent habitual rate. Solid light grey lines represent individual participants. Shaded band represents the 95% confidence interval.

Rate by sentence length interactions

The interactions between rate and sentence length improved the model fit but were not of primary interest. In general, significant interactions demonstrated that rate effects differed in their magnitude across sentence lengths. For simplicity, these results are not reported here, but are included in the model coefficients in Table 23 in Appendix G. Empirical results for each of the sentence lengths are plotted in Figure 34 B. A steeper decline in fast speech is visible for the longer sentences. Note that there is no available data for the 7- and 10-word sentences in the fastest rate (F3), indicating that, following the proportional rate binning procedure, no sentence productions for these items existed at the extreme fast end of the continuum. Sentence productions were balanced in elicitation of the rate conditions but were unbalanced in the proportional rate bins. This was expected.

4.3.2.2 Conversational intelligibility

4.3.2.2.1 Final model

The final model for conversational intelligibility included only the primary variables (group, rate, and their interaction). No other variables were entered into the model. As mentioned above, unlike with previous results, rate was entered as the *rate condition* rather than the proportional rate bins. Therefore, there are six contrasts for this condition (six modified conditions compared to the habitual condition). The fixed effects structure of the final model can be summarized as:

log(100 - Intelligibility) ~ group + rate(condition) +
group x rate(condition)

4.3.2.2.2 Main effects

Conversational intelligibility ratings by group and speech rate are plotted in Figure 36. Model coefficients are reported in Table 24 in Appendix G. The two-way interaction between group and rate was significant but no post-hoc pairwise comparisons are reported for this section.

Speaker group

The same general pattern for speaker group was observed for conversational and sentence reading intelligibility. All group contrasts for conversational intelligibility demonstrated a significant pattern, such that the YC group was rated as most intelligible (YC vs. OC, PD, DBS: $\hat{\beta}$ = -1.607, p <0.001), followed by the OC group (OC vs. PD, DBS: $\hat{\beta}$ = -1.307, p <0.001), with the DBS group rated as least intelligible (PD vs. DBS: $\hat{\beta}$ = -0.793, p <0.001). This was driven by the fact that the YC group demonstrated little change in conversational intelligibility, as is described in the three-way interaction below.

Speech rate

None of the slower rate conditions were associated with differences in intelligibility ratings for conversational speech compared to the habitual rate (S4 vs. H1: $\hat{\beta}$ = 0.085, p = 0.27; S3 vs. H1: $\hat{\beta}$ = -0.044, p = 0.526; S2 vs. H1: $\hat{\beta}$ = 0.03, p = 0.646). Conversely, all three faster rates were associated with significantly worse speech intelligibility ratings compared to habitual speech, as indicated by significant positive estimates for these contrasts (F2 vs. H1: $\hat{\beta}$ = 0.323, p <0.001; F3 vs. H1: $\hat{\beta}$ = 0.392, p <0.001; F4 vs. H1: $\hat{\beta}$ = 0.626, p <0.001).

4.3.2.2.3 Group by rate interaction

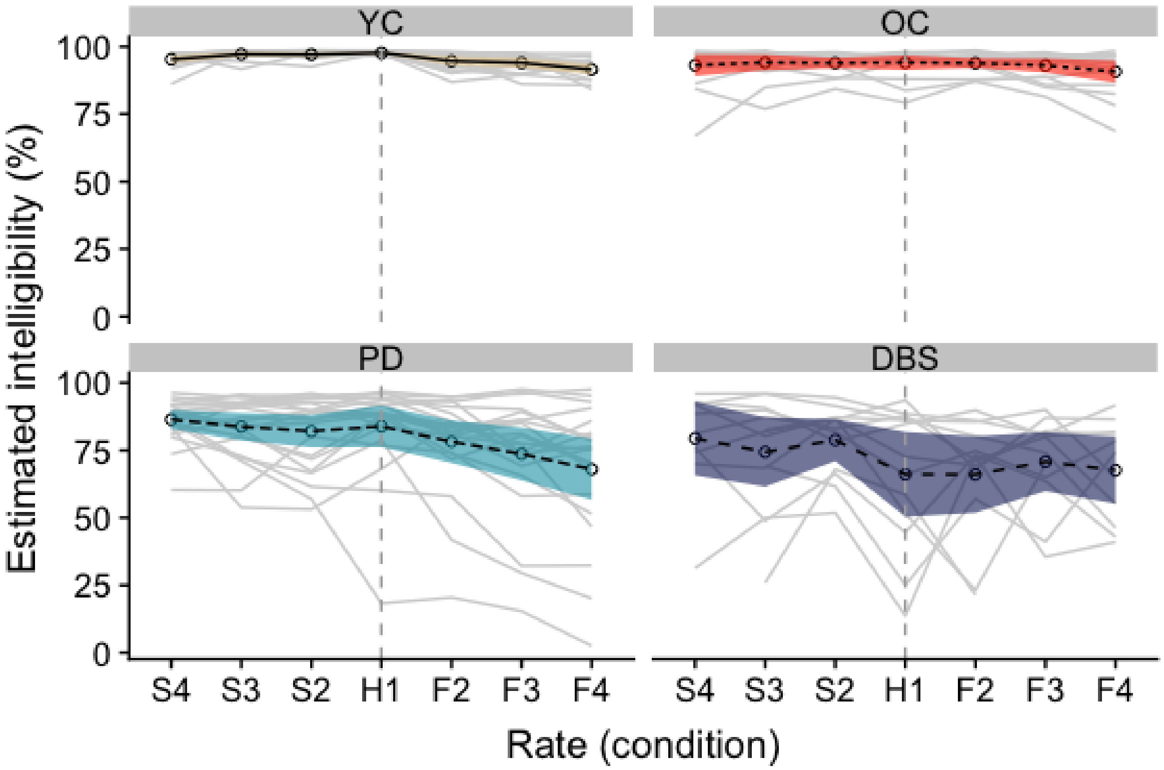


Figure 36: Conversational speech intelligibility by speaker group and rate (condition), averaged over listeners and participants. Solid light grey lines represent individual participants. Vertical line at H1 represent habitual rate. Shaded band represents the 95% confidence interval.

In slow speech, the OC, PD, and DBS groups collectively demonstrated an increase in intelligibility, while the YC groups saw little to no change (in fact, a slight decrease). This was captured by a significant interaction for the YC vs. OC, PD, DBS comparison only at the slowest (S4) speech rate condition ($\hat{\beta}$ = 0.734, p <0.001). The YC vs. Rest contrast was not significant for any other comparisons of slow speech conditions, nor was the OC vs. PD, DBS contrast. The DBS group showed relatively greater intelligibility in some of the slow conditions compared to the PD group, specifically captured by significant interactions for the S2 ($\hat{\beta}$ = 0.511, p = 0.008) and S3 ($\hat{\beta}$ = 0.457, p = 0.024) rates.

In fast speech, the YC group demonstrated a flatter decline than the OC, PD, and DBS groups combined at all fast rates, as seen in the empirical data in Figure 36, though the positive interactions demonstrated a stronger *effect* (F2: $\hat{\beta}$ = 0.512, p = 0.001; F3: $\hat{\beta}$ = 0.587, p <0.001; F4: $\hat{\beta}$ = 0.824, p <0.001).

At the fastest rate, the PD group demonstrated greater decline compared to the DBS group (F4: $\hat{\beta}$ = 0.534, p = 0.028). As with the sentence intelligibility, the empirical data suggested increased intelligibility in the faster conditions for the DBS group, but post-hoc pairwise comparisons indicated these were not significant differences. No other model interactions were significant.

4.3.2.3 Summary of findings of Experiment 3

In summary, in connected speech, represented here by sentence reading and conversational speech, slow speech was generally associated with increases in intelligibility across the groups, while fast speech was associated with decreases. This pattern largely held across all four groups, though the DBS group showed less change in the sentence reading task in slow speech and demonstrated slight *increases* in intelligibility in fast speech; a notable difference compared to the other groups. This trend was also observed in conversation, though to a lesser degree. An expected hierarchy of intelligibility scores was observed across the four groups, with the YC group rated with the highest intelligibility, followed by the OC, PD, and DBS groups as least intelligible. The PD and DBS groups demonstrated much more variability.

While mean intelligibility ratings for the sentence reading and conversational tasks did not demonstrate a significant difference from one another overall, differences did emerge in the group by rate interactions for each of the analyses. Namely, the clinical groups showed relatively less change in intelligibility across rates in the sentence reading task, but more change in conversation.

4.4 Relationship between speech acoustics and intelligibility

As described in Section 3.3.3, in this final analysis, word-level transcription accuracy was modelled as a function of the acoustic variables of interest. This section addresses RQ5 (*what* *is the relationship between speech acoustics and intelligibility?*). Unlike in the previous analyses, model *selection* was not applied; all variables of interest were included in the final model in order to determine how each variable was related to intelligibility when all other variables were considered. These final variables included group and rate, speaker sex, consonant voicing, and the acoustic variables VOT, VDC, QVAI, intensity, and HNR. The variable that captured the presence of audio clipping was also included to account for any impact this may have had on intelligibility. Acoustic variables were not transformed (e.g., VOT was not log-transformed) nor rescaled, but were centered. The variance inflation factor was found to be less than 2 for all acoustic measures, abating concerns of multicollinearity in the model. Visual inspection revealed nonlinear trends for QVAI and HNR, but nonlinear terms did not improve the model fit. The final model thus reflects the linear relationship of all acoustic variables to intelligibility.

The fixed effects structure of the final model can be summarized as:

Proportion words correct (logit-transformed) ~ VOT + VDC + QVAI + Intensity + HNR + voicing + VOT x voicing + VDC x voicing + group + rate + group x rate + sex + clipping

This was an exploratory analysis in which the overall goal was to determine *which acoustic variables had an effect on intelligibility when other factors such as rate and group were controlled for*. As such, the primary goal was to model the main effects of the acoustic variables, and interaction terms *between* acoustic variables and rate were not included[24]. The interaction terms that were included were group x rate, to control for this relationship and to be consistent with the primary research questions, as well as the interaction between consonant voicing and the two stop acoustic measures (VOT and VDC) which are known stop voicing cues (e.g., Davidson, 2016; Lisker & Abramson, 1964). As such, only the interactions related to these acoustic variables will be reported here, though empirical data for both the OC and PD, DBS groups are reported in Figures 37 and 38 for more qualitative

[24] Exploratory analyses did suggest that the main effects of the acoustic variables were similar even when the rate interactions as well as non-linear terms were included in the model.

speculation. Random effects in the model included by-participant random intercepts and slopes for rate (only the S3 and F3 contrasts were included to avoid a singular fit).

It is important to note that this analysis is similar to the intelligibility analyses in Section 4.2, but there are key differences. Specifically, in this analysis, the YC group were excluded and the data were aggregated over consonant PoA and vowel backness. In addition to the primary variables of interest, the acoustic variables, speaker sex and audio clipping were also included. The group and rate variables were included in the model to account for their contribution but are not the primary variables of interest for this analysis.

4.4.1 Main effects

4.4.1.1 Acoustic variables

Acoustic variables that demonstrated a significant effect on intelligibility are displayed in Figures 37 and 38 for the OC and PD/DBS groups, respectively. Model coefficients are reported in Table 25 in Appendix H. Note that the groups are plotted separately for speculative purposes, but the interactions between group and acoustic variables were not included in the model.

The consonant articulation variables, VOT and VDC, did not significantly impact intelligibility on their own (VOT: $\hat{\beta}$ = -1.728, p = 0.155; VDC: $\hat{\beta}$ = -0.103, p = 0.202), though the interaction term between VOT and voicing did ($\hat{\beta}$ = -3.829, p <0.001). There was no significant interaction between VDC and voicing ($\hat{\beta}$ = 0.015, p = 0.804). Panel A in Figures 37 and 38 report this trend. In essence, as VOT becomes longer (e.g., more "voiceless-like"), voiced stops are transcribed with poorer accuracy. Similarly, very short VOT is associated with poorer accuracy as well. VOT does not, for the most part, impact the accuracy of voiceless stops.

Better intelligibility was significantly associated with less vowel centralization (i.e., higher QVAI values: $\hat{\beta}$ = 0.747, p <0.001). Vowel intensity was significantly positively correlated with intelligibility (i.e., higher intensity was associated with higher intelligibility: $\hat{\beta}$ = 0.043, p <0.001). HNR, on the other hand, was significantly *negatively* associated with intelligibility (i.e., better intelligibility was associated with worse voice quality: $\hat{\beta}$ = -0.028, p <0.001).

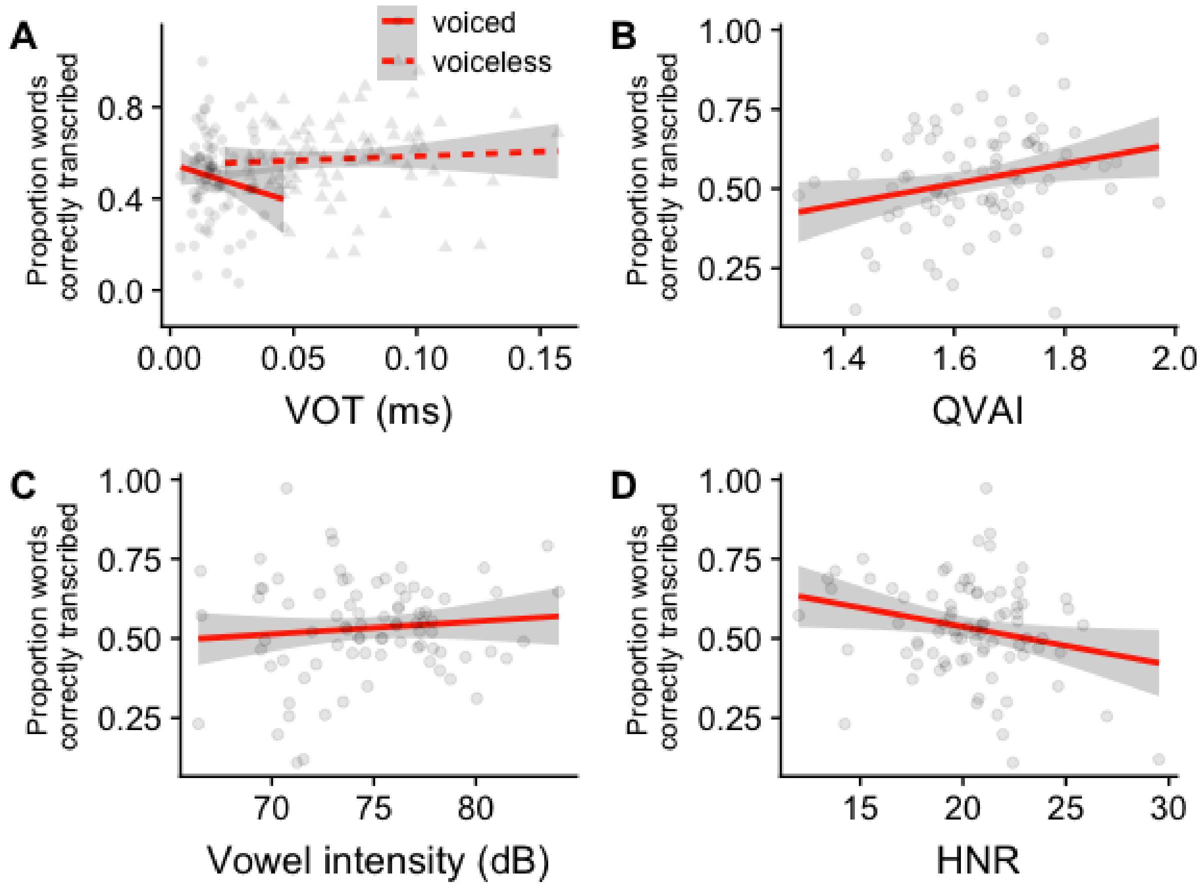


Figure 37: Acoustic variables by intelligibility (proportion words transcribed correctly) for the Older Control group. Individual points represent participant averages for each proportional rate. Only variables that demonstrated significant effects are pictured.

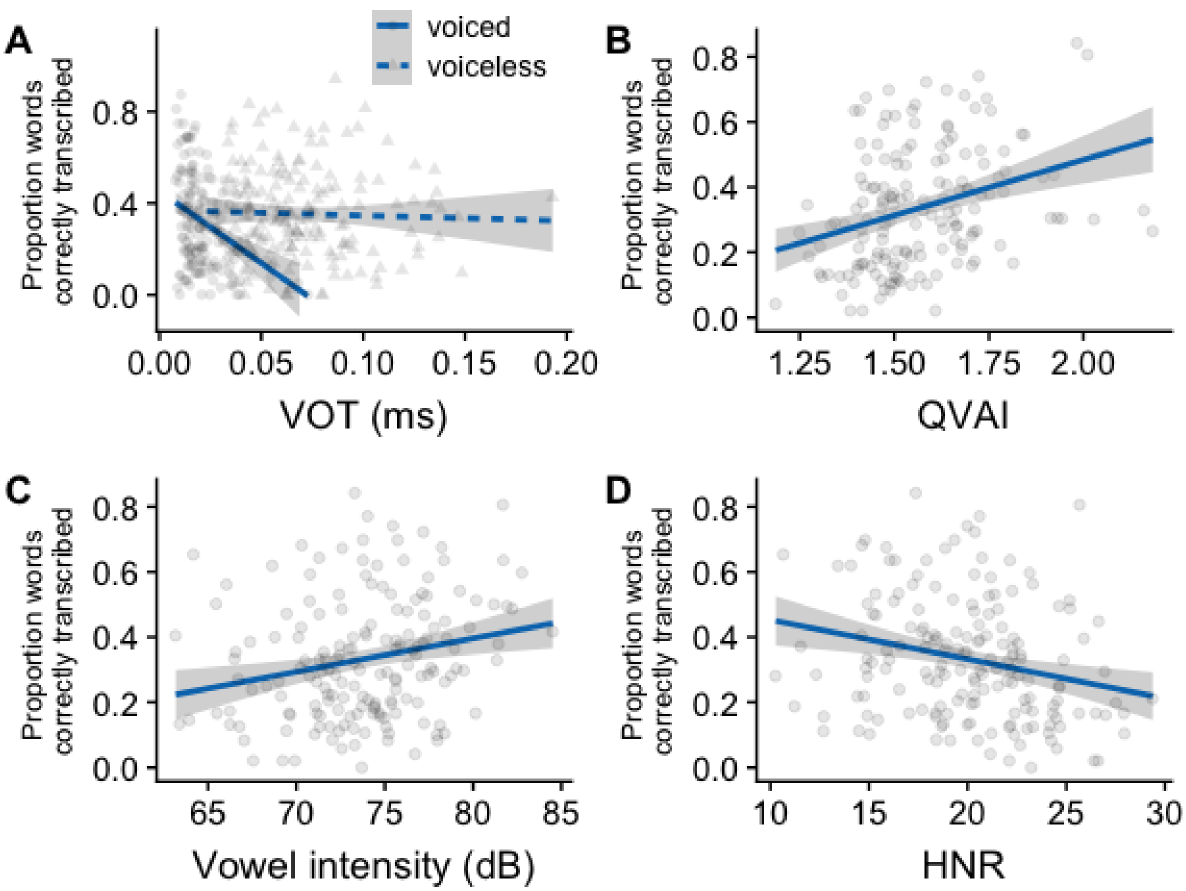


Figure 38: Acoustic variables by intelligibility (proportion words transcribed correctly) for the PD and DBS groups. Individual points represent participant averages for each proportional rate. Only variables that demonstrated significant effects are pictured.

4.4.1.2 Other variables

As mentioned, group, rate, and their interaction were included in the model in order to account for changes in intelligibility across these measures (as were reported in Section 4.2). It is worth noting that, as one would expect, in the current model of word-level intelligibility, their effects patterned with what was seen in the consonant and vowel results. That is, the OC group was transcribed with greater accuracy than the PD and DBS groups, ($\hat{\beta}$ = 0.358, p = 0.002), but the two clinical groups did not differ ($\hat{\beta}$ = 0.003, p = 0.979). Similarly, compared to habitual speech, slower speech was transcribed with lower accuracy (S3 vs. H1: $\hat{\beta}$ = -0.182, p = 0.002; S2 vs. H1: $\hat{\beta}$ = -0.186, p <0.001), but fast speech did not differ (F2 vs. H1: $\hat{\beta}$ = -0.073, p = 0.115; F3 vs. H1: $\hat{\beta}$ = 0.003, p = 0.965).

Female speakers were transcribed with greater accuracy than male speakers ($\hat{\beta}$ = 0.161, p = 0.003). Audio clipping was not found to influence speech intelligibility ($\hat{\beta}$ = -0.115, p = 0.313).

In summary, of the acoustic variables studied in this thesis, VOT as an index of voicing, vowel centralization, speech intensity, and voice quality were all found to impart a significant effect on word-level speech intelligibility in a group of older healthy controls and individuals with PD with and without DBS.

5 Discussion

This chapter begins with a restatement of the original research questions and an overview of the findings. Each research question is addressed in the context of existing relevant literature, and future directions and clinical implications are discussed at the end. Limitations to the present study not mentioned earlier are addressed in Section 5.6.

The overall aim of this study was to investigate the changes that occur in spoken communication in PD along a wide continuum of self-selected speech rate adjustments. In particular, this study addressed a set of acoustic characteristics that were hypothesized to be a function of speech rate modifications for individuals with PD and hypokinetic dysarthria (HkD), as well as individuals who may have additional speech symptoms following DBS, a common surgical intervention. These findings are compared primarily with older neurologically healthy controls. This study also explored how modified speech rates impacted speech intelligibility across multiple speech tasks and compared the three aforementioned groups alongside a group of younger healthy control speakers (for a subset of comparisons). Slower speech is a frequently recommended treatment target for some individuals with PD and HkD, but recent literature has suggested that slower speech is not always associated with gains in intelligibility for these individuals. This study probed a wider range of speech rates and speech tasks than those examined in previous studies in order to better understand how speech rate modulates aspects of speech production that impact a speaker's likelihood of being understood.

5.1 Overview of research questions and main findings

As stated in Section 2.6.1, the primary research questions were as follows. Hypotheses will be discussed in the context of the findings in greater detail in Sections 5.2 – 5.5.

1. What differences in terms of the range of self-selected speech rates exist across speaker groups (younger and older controls, people with PD with and without DBS) when instructed to modify their rate from very slow to very fast?

2. What are the acoustic-phonetic changes that occur in PD and control groups along a speech rate continuum?

3. How does such a continuum of speech rates impact speech intelligibility in PD and control groups?

4. What differences in speech intelligibility exist across speech tasks along a speech rate continuum?

5. What is the relationship between speech acoustics and intelligibility within a speech rate continuum?

Overall, the main findings of the present study were as follows:

1. Individuals with PD and DBS, as well as younger and older healthy control groups, successfully modulated their speech rate along a wide range from very slow to very fast. The PD and DBS groups had a similar range to that of controls for slow rates, but a relatively more restricted range at fast rates. This was particularly the case for individuals with DBS.

2. While all groups demonstrated similar changes to consonant and vowel articulation in expected directions along the rate continuum, individuals with PD and DBS demonstrated longer voiced VOT, more voicing through stop closure, and more vowel centralization overall. The PD groups also did not demonstrate expected changes in consonant distinctiveness along the rate continuum, whereas this pattern was observed for the controls.

3. Slower speech was associated with lower speech intensity for all groups compared to their habitual rates. Faster speech was associated with higher speech intensity than slower speech. Voice quality was inversely related to speech rate for all but the DBS groups; specifically, higher voice quality was observed in slow speech, and lower voice quality in fast speech. The DBS group demonstrated the opposite trend.

4. An asymmetry was present between phoneme identification and sentence estimation. Phoneme identification was overall lower in both fast and slow speech compared to speaker’s habitual rates, whereas estimated sentence and conversational intelligibility ratings were higher in slow speech and lower in fast speech overall.

5. Voiced VOT, vowel centralization, speech intensity, and voice quality were all associated with speech intelligibility (word accuracy). All of these but voice quality were in the expected directions, such that longer voiced VOT, more centralized vowels, and quieter speech were all associated with poorer intelligibility. Conversely, poorer voice quality was associated with higher intelligibility.

5.2 RQ1: Group differences in self-selected speech rate modifications

Consideration of rate modification differences will follow a discussion of the differences in habitual rates observed in the four groups.

5.2.1 Habitual rate differences

The present study specifically was designed to include individuals with PD (with and without DBS) who had documented speech changes affecting articulation. Many of these individuals also experienced changes in speech rate. A comparison of habitual speech rate among the four groups in Section 4.1.1.1 indicated that the PD and DBS groups did indeed have faster speech rates compared to the older healthy controls (PD: 167 WPM; DBS: 167 WPM; OC: 147 WPM), though the difference between the OC and DBS groups did not reach significance. Conversely, the habitual rate for the clinical groups was nearly identical to the younger healthy control group mean (169 WPM).

Slowed speech in healthy older adults compared to younger talkers is a consistent finding in the literature and has been observed across a multitude of speech tasks (e.g., Jacewicz et al., 2009; Fletcher et al., 2015; Liss, 1990; Smith, Wasowicz, & Preston, 1987; Wohlert & Smith, 1998). The present study explored habitual speech rate in a simple sentence reading task (“Please say ___ again”). Mefferd & Corder (2014) suggested that a slowed habitual speaking rate in older adults may be due to a compensatory strategy in the face of reduced articulatory stiffness. Previous literature has also documented cognitive-linguistic decline in older adults (Bryan, Luszcz, & Crawford, 1997; Burke & MacKay, 1997; Elgamal, Roy, & Sharratt, 2011; Glosser & Deser, 1992; Lamar, Resnick, & Zonderman, 2003) which may also be implicated in a slowed speaking rate to some degree (Nip & Green, 2013).

Conversely, faster speech in individuals with PD is sometimes noted. Faster speech, and specifically "short rushes of speech" was a deviant speech characteristic noted in the early work of Darley, Aronson, and Brown (1969a; 1969b; 1975), but many studies since have found unimpaired or slower speech rates at the group level in PD (Connor et al., 1989; Hsu et al., 2017; Kleinow et al., 2001; Ludlow et al., 1987; Martínez-Sánchez et al., 2016; Skodda & Schlegel, 2008; Tjaden & Wilding, 2004; Walsh & Smith, 2012; Weismer et al., 2001). It is likely that people with PD who demonstrate these patterns of faster speech represent a distinct phenotype (Adams & Dykstra, 2009). The PD speakers in the present study were recruited on the basis that there was *some* mention of speech disturbances relating to articulation in their charts (e.g., "mumbling", "tachyphemia", "slurred"). The DBS group, on the other hand, were recruited or referred for the study in many cases without knowledge of their speech deficits. Thus, while the PD group specifically represents a group of individuals with PD and speech deficits, the DBS group may better represent the speech symptoms of the DBS population at large.

Tsuboi et al. (2014) reported on patterns of observed speech deficits in people with PD and DBS and found five distinct clusters. Approximately one quarter of their sample demonstrated relatively unimpaired speech, while another quarter demonstrated speech rate abnormalities and disfluencies. Three clusters accounting for the remaining 50% included breathy type, strained voice type, and spastic dysarthria type. While the DBS group in this study did not demonstrate a statistically significant difference in habitual speech rate from the older adults, they did show nearly identical habitual speech rates to the PD group, who were found to have a faster habitual rate. This discrepancy may have been related to the higher degree of variability (and smaller sample size) in the DBS group.

5.2.2 Modified speech rate range differences

Despite a finding of faster habitual rates, both PD and DBS speaker groups were successful in modifying their speech rate along a continuum from slow to fast. In essence, they were able to slow their speech to a similar degree as the healthy older talkers but did not increase their speech rate to quite the same extent. These findings partially support the original hypothesis: both control groups produced a wider range of *fast* rates compared to the clinical groups, but, overall, did not differ at slow rates.

Early work suggested that individuals with PD were unable to voluntarily modify their speech rates, originally thought to be related to muscular rigidity (Ludlow & Bassich, 1984). More recent studies have demonstrated that people with PD are able to modify their speech rate (Martens et al., 2015; McRae et al., 2002; Tjaden, 2000a, 2003; Van Nuffelen et al., 2010, 2009), but in some cases to a lesser degree than control speakers (Kleinow et al., 2001). Difficulties in maintaining or manipulating speech rate is hypothesized to be related to difficulties with sensorimotor integration (Forrest, Nygaard, Pisoni, & Siemers, 1998; Millian-Morell et al., 2018).

In a novel study in which PDs as well as younger and older healthy talkers were asked to produce a phrase at different rates and then later transcribe their own recordings, Forrest et al. (1998) found asymmetries in individuals with faster versus slower habitual rates of speech. Specifically, individuals with PD who demonstrated faster habitual rates of speech transcribed their own speech with poorer accuracy when it was played back to them. The findings of Forrest et al. (1998) led the authors to suggest that speech rate abnormalities in PD may be related to difficulties in perceiving spoken language, for example relating to important temporal differences in speech such as lexical stress. Forrest et al. (1998) also proposed that sensory deficits, which are documented in other domains such as tactile, visual, and auditory (Artieda, Pastor, Lacruz, & Obeso, 1992), may also be implicated in speech movement control. Perceptual deficits have been further reported in respiratory sensation (Hegland, Troche, & Brandimore, 2019), vocal emotions (Breitenstein, Van Lancker, Daum, & Waters, 2001), and speech intensity (Clark et al., 2014).

With regards to slow speech, both clinical groups slowed their speech to a similar degree compared to the older and younger controls. In fact, the older controls had a slightly narrower slow speech range (98 WPM versus 107 and 105 for the PD and DBS groups, respectively, though this difference did not reach statistical significance). The younger group was found to have a wider range of slow speech rates (a mean difference of 122 WPM slower than their habitual rates), though this difference only reached significance when compared to the older healthy talkers, not the clinical groups.

Differences among the control and clinical groups did emerge in the faster speech conditions. Specifically, the YC group produced the widest range of fast speech (an average of 185

WPM faster than their habitual rates), followed by the OC group (138 WPM faster). Both the PD and DBS groups demonstrated a narrower range of fast speech (113 and 90 WPM, respectively), though for the clinical groups these differences only reached significance for the OC versus DBS comparison.

Tsao and colleagues investigated the characterization of speech rate adjustments in healthy talkers and suggested that speech rate ranges are modulated by neuromuscular, rather than sociolinguistic control (Tsao & Weismer, 1997; Tsao et al., 2006). This proposal was driven by their findings that, regardless of habitual rates, habitually “fast” or “slow” talkers adjusted their rate when asked to do so to a similar degree; i.e., speakers had distinct intercepts but identical slopes of change. Even when asked to speak at a rate that felt maximally fast, individuals produced similar proportional changes to their habitual rates (Tsao & Weismer, 1997). In the present study, treating speech rate as a proportion of each talker’s habitual rate was intended to take this into account. Therefore, the “slowest” and “fastest” rates reported in this study are indeed designed to consider “slowest” and “fastest” specific to an individual.

Tsao and colleagues as well as others also found that healthy talkers modify fast and slow rates in a nonlinear manner Specifically, talkers made larger adjustments in slow speech, and smaller adjustments in fast speech (Adams, 1993; Tsao & Weismer, 1997; Tsao et al., 2006). This was captured by finding steeper and flatter slopes of change in slow and fast speech, respectively. This pattern was replicated in the present study, as can be seen in Figure 5 in Section 4.1.1.2.

In summary, PD and DBS groups were found to have faster habitual rates than older healthy controls and were able to increase and decrease their rates when instructed to do so. They were able to produce slow rates to a similar degree to the healthy controls but demonstrated a more restricted range on the fast ends of the continuum.

5.3 RQ2: Group differences in acoustic changes along a speech rate continuum

Group differences in RQ2 were addressed in Experiment 1 using the two-way interactions involving group and rate, the three-way interactions involving group, rate, and additional linguistic variables (e.g., consonant voicing), as well as the subsequent post-hoc pairwise comparisons to explore changes across specific rates. Acoustic *distinctiveness* can be inferred from the interactions between speech rate and the phonological variables such as consonant voicing or vowel place. The main effects of rate, as well as interactions with rate that did not involve group essentially answered the question "across these groups, what acoustic changes are observed as a function of rate?"

Acoustic variables of interest included stop consonant variables (VOT, VDC), vowel centralization (QVAI), speech intensity, and voice quality (HNR). All acoustic measures were derived from the nonsense word sentence task containing stop consonants. The acoustic analyses focused exclusively on the OC, PD, and DBS groups (i.e., not the YC group).

Overall, the main results can be summarized in the following way. Of the five acoustic variables studied, only HNR demonstrated a two-way interaction between group and rate, indicating that, in the absence of other mediating factors, the three groups adjusted their speech in similar ways along a rate continuum. Significant three-way interactions involving group and rate were observed for VOT with stop voicing, speech intensity with speaker sex, and HNR with vowel height, indicating that group differences in speech changes along a rate continuum were modulated with respect to another linguistic or talker-specific variable.

Acoustic measures are discussed separately in terms of consonant and vowel acoustic measures (VOT, VDC, QVAI), which reflect laryngeal and supralaryngeal vocal tract adjustments impacting phonemic categories, and voice acoustic measures (intensity, HNR), which reflect laryngeal and respiratory adjustments impacting overall voice production.

5.3.1 Articulatory acoustics

Answering the question of group differences in the context of acoustic distinctiveness, that is, the extent to which speakers maintained, increased, or decreased the phonological contrast between consonants or vowels by way of these adjustments, involves examining interactions with the linguistic variables. For the stop consonant measures (VOT[25] and VDC),

---

[25] Note that Auzou et al. (2000) suggested using the terms "short-lag" and "long-lag" to refer to "voiced" and "voiceless" stops to better characterize these distinctions in dysarthric speakers across languages. While this point is sensible, the more common "voiced"/"voiceless" terminology will be used in this section to be consistent with the majority of the literature on VOT.

phonological categories of voicing and place of articulation were also considered. QVAI is a composite variable that considers acoustic measures of all four corner vowels in a single measure. No additional phonological variables were included for this analysis. An increase in QVAI can be inferred as an increase in vowel distinctiveness.

5.3.1.1 Stop consonant acoustics

5.3.1.1.1 Voice onset time (VOT)

An important finding was the significant *three-way* interaction between group, rate, and consonant voicing for the VOT analysis. This interaction demonstrated that the healthy older talkers *increased* the voiced-voiceless VOT contrast in slow speech (specifically, at their slowest speech rates), and *decreased* this contrast (i.e., more overlap) in fast speech, consistent with hypotheses. The PD and DBS groups, on the other hand, demonstrated a similar degree of stop voicing distinctiveness across the rate continuum. That is, speech rate did not impact their consonant distinctiveness (in terms of VOT) as much in either direction. Furthermore, there were no observed statistical differences between the DBS and PD groups for this interaction, despite the finding that DBS participants demonstrated longer overall VOTs compared to the other groups. That is, while differences across groups were observed for VOT, VOT distinctiveness was only affected by speech rate for the healthy older group.

Despite the fact that VOT and VDC are both temporal measures of speech production that are known to be sensitive to dysarthria (Kent et al., 1999; Weismer et al., 2012), neither have previously been the focus of speech rate manipulation studies in clinical populations.

In young healthy talkers, previous literature has demonstrated that increased voiced and voiceless VOT distinctions vary by speech rate, and that this is largely driven by changes to voiceless VOT production (Miller et al., 1986). Voiced VOTs, on the other hand, are known to vary much less with speech rate compared to voiceless VOTs (Kessinger & Blumstein, 1997; Miller et al., 1986; Miller, O'Rourke, & Volaitis, 1997; Summerfield, 1981). The findings here support this for the neurologically healthy geriatric speakers but suggest that talkers with HkD modulate VOT (particularly voiced VOT) to different extents.

Overall, the PD and DBS talkers in this study demonstrated more changes in *voiced* versus voiceless VOT. This can be seen as a steeper rise for voiced VOT in slow speech in the right-hand panel of Figure 12, and a greater quantity of significant pairwise comparisons for voiced stops. One interpretation could be that the reason that PD and DBS talkers did not increase consonant distinctiveness as much as the controls did in slow speech was not due to expected increases in voiceless VOTs, but to increases of an *unexpected greater magnitude* for voiced VOTs. Statistically, this was found to be the case for the DBS group. This can be seen in the pairwise comparisons (Figure 14). In most cases, and particularly in slow speech, voiceless VOTs did not differ across the three groups, whereas voiced VOT did. Voiced VOTs were longer for the DBS group at both the slowest (compared to the OC group) and slower (compared to both OC and PD groups) rates.

The existing literature on VOT in PD does not point to consistent trends. The present study corroborated findings of *longer* VOT, overall, in individuals with PD, and longest VOTs in those with DBS. This was mainly due to longer voiced rather than voiceless VOTs, as evidenced by the pairwise comparisons (Figure 14). The finding of similar voiceless VOT in PDs and controls here is consistent with the majority of studies that have explored differences in voiceless stop production in PD (Bunton & Weismer, 2002; Connor et al., 1989; Cushnie-Sparrow et al., 2016; Fischer & Goberman, 2010; Forrest et al., 1989; Ravizza, 2003; cf. Flint et al., 1992).

Some studies have additionally reported longer voiced VOT (Forrest et al., 1989) in talkers with PD, making them more voiceless-like in nature, while others have reported more overlap between categories in general (Lieberman et al., 1992; Miller et al., 1986). These findings are supported by the present study across the rate continuum for the DBS speakers, and perhaps especially so in slow speech. While abnormal VOT is a considered to be a reflection of difficulties in coordinating the laryngeal and supralaryngeal system (Weismer, 1984a), Auzou et al. (2000) suggested that abnormal VOT in PD may also be attributable to abnormal lung volume during speech.

One previous study did explore VOT in a graded speech rate task in talkers with PD (Tjaden, 2000a). VOT itself was not a primary outcome measure except in the habitual condition. There was a trend of longer voiced VOTs in the PD group, despite shorter vowel durations (a proxy of faster habitual speech rates). Importantly, the author found that VOT was not associated with coarticulatory formant patterns in fast or slow speech for either group. That is, while the variation in VOT along speech rate itself was not studied, VOT was not predictive of other coarticulatory speech rate changes.

VOT production in DBS has received relatively little attention. There are reports of shorter VOTs produced by people with DBS in alternate motion speech tasks (Putzer et al., 2008), while others have found no differences (Karlsson et al., 2012). Chenausky et al. (2011) reported increased VOT *variability* in talkers with DBS compared to healthy controls but did not directly report on VOT itself. Increased variability is a common pattern throughout the present study with regards to the DBS findings.

While not directly related to VOT, spirantization (i.e., incomplete stop closure allowing for a leakage of air making the stop more fricative-like) is a measure that has been reported with greater frequency in individuals with DBS (Chenausky et al., 2011; Dromey & Bjarnason, 2011; Eklund et al., 2014; Karlsson et al., 2014). Karlsson et al. (2014) found that individuals with DBS of the subthalamic nucleus or caudal zone incerta exhibited greater degrees of spirantization compared to their preoperative speech and when DBS was off during passage reading. Interestingly, however, they also found that these talkers produced more prominent stop releases, attributable to a *stronger* stop occlusion. These findings would appear to contradict one another, but Karlsson et al. (2014) suggested that while these individuals were able to generate sufficient energy during speech to produce a distinctive plosive release (compared to when DBS was off), a consequence of this was premature stop consonant frication. Many of the stops in the current data were noted as having spirantization, but this was not categorically measured. Future work should explore the relationship between VOT, spirantization, and spectral stop moments and intensity in PD, and especially in those with DBS.

Related to spirantization, another consideration is the presence of VOTs that could not be measured. In the current study, stops that had a clear release throughout the entire stop but had no clear closure were coded as having measurable VOT, but in these instances the duration of VOT was equal to that of the entire stop consonant (i.e., no or little closure). These would likely be cases of longer VOT. These instances were kept for the VOT analysis but removed from the VDC analysis. Karlsson, Unger, et al. (2011) considered these cases a form of unmeasurable VOT, and noted that these types of instances were more common in individuals with DBS. While these extreme cases were uncommon in the present data, accounting for less than 3% of the data, they were more common in individuals with PD and DBS (OC: <1%; PD: 2.5%; DBS: 3.4%). Exploratory analyses with these data points removed did not change the results. The proportion of VOT to closure duration would be another metric worth considering in order to explore these potential effects.

Another type of aberrant VOT production is “unreleased” stops, i.e., stops produced with no obvious burst. These were excluded from the VOT analysis. In the present study, 3.7% of stops overall were unreleased. This amounted to 1.86% for controls, 3.85% for PDs, and 7.81% for the DBS group. This demonstrates a similar but attenuated pattern to that reported by Özsancak et al. (2001), in which 19% of stops in talkers with HkD could not be measured due to the absence of a clear burst, compared to 7% in controls. Exploring more measures of aberrant stop production in combination with VOT measures may be a promising avenue for determining underlying acoustic and physiological underpinnings of differences in laryngeal-supralaryngeal coordination impairments in HkD.

5.3.1.1.2 Voicing during closure (VDC)

There was no difference in VDC between the groups, nor were there any observed interactions between group and speech rate in the present study, contrary to predictions. The OC group did maintain a stronger distinction between voiced and voiceless stops however, indicated by a lower production of total VDC in voiceless stops compared to the clinical groups. This was captured by the group by voicing contrast, and is visible in Figure 16 A. The DBS group displayed the least amount of contrast, evidenced by the interaction for the PD vs. DBS comparison. This was demonstrated by a significant interaction between group and voicing. A trend that was not borne out by the statistical analyses, but which is visible in Figure 17, is that the DBS speaker group appeared to demonstrate an *especially* steep increase in VDC in voiceless stops in faster speech.

As with VOT, VDC has received very little attention in the literature on speech rate modifications. Weismer (1984b) explored VDC in a group of people with PD as well as younger and older healthy control groups, who were asked to produce sentences at a faster-than-normal rate. In the Weismer (1984b) study, VDC was split on the basis of “voicing [occurring] for more than 20ms into voiceless stops[26]", which differs from the dichotomization in the present study. Weismer (1984b) found that while there was a slight trend for more voicing during closure of voiceless stops in PDs compared to healthy older adults, this was not a clearly distinguishing feature. A clear difference was, however, between younger and older speakers, the former almost never producing more than 20ms of voicing into voiceless stop closure. The author thus attributed the increased presence of VDC to normal muscular deterioration of laryngeal tissue, suggested that longer periods of VDC might reflect difficulty in initiating the laryngeal devoicing gesture as a result of these structural changes.

With this in mind, it is worth mentioning that the present study did not examine acoustic productions of the younger speakers. Doing so would shed light on the relative contributions of aging biomechanics and rate. Previous literature has also suggested that VDC in voiceless stops is not uncommon when following a sonorant (Davidson, 2018), and VDC in general is highly dependent on the surrounding phonetic environment in general (Davidson, 2016).

Another study reported no differences in "voicing intrusion" errors (i.e., VDC) in PD subjects versus older controls, but did report that subjects with clinically diagnosed depression demonstrated consistently more VDC compared to controls, and slightly more compared to PDs (Flint et al., 1992). Symptoms of depression may appear similar to those of early PD in terms of slowness of movement and speech disturbances (Flint et al., 1992; Lohr & Wisniewski, 1987). Depression is also a common co-occurrence in PD (Reijnders, Ehrt, Weber, Aarsland, & Leentjens, 2008). While depression was not controlled for in the present study, should be considered as a possible factor. Flint et al. (1992) found that the most consistent acoustic measure differentiating the PDs from individuals with depression was speech rate, such that individuals with PD demonstrated faster rates of speech.

[26] Exploratory analyses in the present study suggested that binning the data in this way led to similar results (i.e., three-way interaction did not improve fit). Treating VDC as a continuous (logit-transformed) variable, however, was associated with a three-way interaction between group, rate, and voicing. Because of the high number of stops with total voicing through closure in the data, however, the current binning procedure was elected as the best approach, but future studies should consider differences in conclusions based on treatment of proportional variables such as VDC.

While not a direct measure of coarticulation, the finding of more VDC in voiceless stops for the PD group overall, coupled with the finding that there is overall more VDC present in fast speech, may be cautiously interpreted as evidence of "blurred" acoustic contrasts in the parkinsonian speech (Kent & Rosenbek, 1982; Tjaden, 2000b; Weismer, 1984a). That is, the PD group demonstrated speech production patterns more likely to occur in fast speech (more VDC in voiceless stops).

An alternative interpretation of this finding could also be that the PD and DBS groups, who were found to indeed produce faster rates in this task compared to the healthy older controls, simply produced more VDC *because they were faster*. The rate metric in this study critically explored *proportional* rates of speech to control for this, but a closer look at absolute rates of speech across talkers would give more insight into this pattern. A more in-depth comparison between the clinical groups and the *younger* healthy control group, who also produced a faster rate of speech, could shed light on whether these differences are related to speech rate or aging, or both.

VDC was also found to be produced with more overlap across articulatory place categories at the extreme ends of the continuum, as evidenced by a significant interaction between rate and place of articulation for VDC at the slowest and fastest rate comparisons. Tjaden and colleagues found that spectral stop differences between /t/ and /k/, measuring place of articulation distinctiveness, were smaller for individuals with PD but did not demonstrate further changes in slow speech for most speakers (Tjaden & Wilding, 2004). The current study did not demonstrate a three-way interaction between rate, place, and group, however. A within-speaker approach would facilitate our understanding these relationships better in future work (Feenaughty et al., 2014; Yunusova et al., 2005).

It is worth mentioning that this study only measured positive VOT and that negative VOT was not considered, and in fact hardly seen. This was based on a definition in which negative VOT would have had to involve voicing starting during the closure, prior to VOT (i.e., prevoicing). This criterion did not operationally define stops that had continuous voicing through closure as having negative VOT (i.e., voicing "bleed"; Davidson, 2016). Including stops with complete voicing through closure as having negative VOT could alter the pattern of results and should be considered as a point of comparison in the future. Further systematic examination of the frequency of prevoicing, even if small, could have important implications as well.

In summary, in terms of consonant *distinctiveness*, the clinical group produced less contrast overall as evidenced by longer voiced VOTs and more VDC in voiceless stop closures. Speech rate demonstrated a clear effect on stop distinctiveness for the healthy geriatric group's VOT production, in that they produced greater contrast in slow speech and less in fast speech. VDC as a metric of *distinctiveness* was not affected by rate for any of the groups. Speech rate did not affect the PD groups' acoustic distinctiveness as it did for the controls. That is, while the PDs predictably produced longer VOT and less VDC in slow speech and the inverse in fast speech, the difference between voiced and voiceless VOTs did not change in the same way that it did for the healthy talkers.

5.3.1.2 Vowel acoustics

With regards to vowel articulation, controls produced larger vowel spaces (larger QVAI) compared to the PD groups, and the DBS group exhibited the most centralized vowel spaces (smallest QVAI). As predicted, all three groups demonstrated similar degrees of vowel expansion (larger QVAI) in slow speech and more centralization in fast speech, as evidenced by the rate effect and lack of an interaction between group and rate. This rate effect is consistent with studies of rate modification that have looked at vowel expansion in talkers with PD (Buccheri, 2013; McRae et al., 2002; Tjaden et al., 2005; Tjaden & Wilding, 2004) and in neurologically healthy talkers (Fletcher et al., 2015; Fourakis, 1991; Lindblom, 1963; Tjaden & Wilding, 2004; Tsao & Iqbal, 2006; Tsao & Weismer, 1997; Tsao et al., 2006; Turner et al., 1995; Weismer et al., 2000).

The overall vowel articulation patterns observed in the present study largely support the current literature. Specifically, talkers with HkD demonstrated smaller vowel spaces compared to controls (Lam & Tjaden, 2016; Lansford & Liss, 2014; McRae et al., 2002; Rusz, Cmejla, et al., 2013; Skodda et al., 2012, 2011; Tjaden et al., 2013a; Watson & Munson, 2008; Whitfield & Goberman, 2014), and more so for talkers with DBS (Sidtis et al., 2016; cf. Tanaka et al., 2016), and female talkers had larger vowel spaces than males (Byrd, 1994; Fletcher et al., 2017a; Jacewicz et al., 2009; Neel, 2008).

One interesting finding was the interaction between speech rate and speaker sex, demonstrating that female talkers demonstrated more extreme effects at both ends of the rate continuum. That is, they showed greater vowel expansion in slow speech, and greater reduction in fast speech. Females also demonstrated a larger vowel space overall, which is consistent with the literature that suggests this is due to a combination of sociolinguistic and biomechanical factors (Fant, 1966, 1970, 1975; Henton, 1995). It should be noted that all sex-specific differences in the present study should be considered with caution, however, given the imbalance of male/female participants across groups (i.e., there were only four females in the PD group and two in the DBS group, whereas the control groups had a nearly even split). That being said, there were no interactions between sex and group (or sex, group, and rate) for the vowel measure QVAI, so sex differences that were found correspond to values averaged over the groups.

The finding that vowel expansion was associated with slow speech, and females tended to exhibit slower habitual rates of speech than men, could partially explain this finding on the *slow* end of the continuum if it were the case that females were simply slowing down more. This would not, however, explain differences in the opposite direction, in which female talkers showed even more centralization in fast speech. Visual inspection of the data suggested that females produced overall slower speech (longer utterance durations) in habitual but produced a wider range of proportional rates (slower and faster extremes).

Tsao and Iqbal (2006) found differences in vowel expansion for habitually fast and slow male and female talkers. The authors found that overall, larger vowel spaces were associated with female talkers and slower talkers, though these groups also demonstrated substantially higher amounts of variability and overlap across groups. These data were part of larger study that explored a speech rate continuum in healthy talkers and found that, regardless of whether they were habitually “fast” or “slow” talkers, males and females manipulated their speech rate to the same degree (Tsao et al., 2006). That is, fast and slower talkers demonstrated distinct intercepts or “launch points” for rate adjustments but adjusted their rates in the same way (i.e., identical slopes). Differences between male and female talkers may also be mediated to some degree by sociolinguistic factors and vocal tract size differences (Jacewicz et al., 2009; Simpson & Ericsdotter, 2007).

The magnitude of vowel adjustments was not as large for slow speech, which could be attributable to the task. While this was a connected speech task, it was fairly contrived and contained novel words. It is likely that individuals may have been hyperarticulating their speech even at their habitual rates, more than they would have for a spontaneous speech task, resulting in less noticeable impairment (Bunton & Keintz, 2008; Ho, Iansek, & Bradshaw, 2002; Sidtis et al., 2012, 2010). Less common words are also known to be produced with greater vowel space than high-frequency words for individuals with and without PD (Munson & Solomon, 2004; wright2004; Watson & Munson, 2008). Future extensions of this work should explore changes in the vowel production of the spontaneous speech samples.

Taken together with the VOT and VDC findings, acoustic distinctiveness was maintained across the rate continuum for healthy speakers for both vowels and consonants, but this was only the case for vowels for the PD groups. It was not the case that the PD groups' consonants became *less* distinct when they modulated their rate, but rather, they maintained a degree of contrastiveness that was already reduced compared to the healthy talkers.

5.3.2 Voice acoustics

There were no main group differences of either vowel intensity or voice quality. Regarding the primary research question (*what are the acoustic changes across groups and speech rates?*), the DBS group demonstrated marked differences in how they altered their voice production. DBS talkers showed marked clear differences in terms of the *degree* to which they got quieter in slow speech, and the degree and direction of their voice quality adjustments across the rate continuum. Specifically, the OC and PD groups were *quieter* and had *the same or better voice quality* in slow speech. The DBS group was also quieter in slow speech, but to a greater degree. The DBS group also demonstrated *worse* voice quality in slow speech. These differences are discussed in the following two sections.

5.3.2.1 Intensity

With regards to vowel intensity, there was a three-way interaction between group, rate, and speaker sex. Healthy older males demonstrated a greater decrease in speech intensity at slow rates than did the PD and DBS male talkers, accounting for this interaction.

There were no overall group effects on intensity, indicating the clinical participants did not exhibit lower speech intensity as measured on the vowel. Lower speech intensity, also known as *hypophonia* (Duffy, 2013) is one of the most common speech symptoms associated with PD (Adams & Dykstra, 2009). While this is borne out in several acoustic studies comparing people with PD with age-matched controls (Adams, Haralabous, Dykstra, Abrams, & Jog, 2005; Fox & Ramig, 1997; Ho et al., 1999; Tjaden et al., 2013a), others have documented similar intensity levels in groups of individuals with PD compared to healthy controls (Canter, 1963; Ludlow & Bassich, 1984; Metter & Hanson, 1986). It should be noted that speech intensity is often measured at the phrase level, whereas in this study it was measured on the vowel. The individuals recruited in this study were not recruited specifically for exhibiting hypophonia, but rather articulatory speech impairments. It is also important to note that intensity was measured on the vowel and not across the whole utterance. It is possible that differences would emerge if sentence or breath group intensity had been measured.

While the clinical groups did not demonstrate overall reduced vowel intensity, an interesting finding was that across all three groups in the present acoustic study *reduced* their speech loudness in slow speech. This is not entirely consistent with studies that have, for example, investigated loud versus slow speaking conditions (Tjaden & Wilding, 2011c, 2004) or louder speech following rate reduction from DAF (Hanson & Metter, 1983).

Tjaden and Wilding (2011c) reported that habitual speech for speakers with PD or MS and controls was significantly louder in slow speech, but that this was attributable only to a 1 dB difference. These differences were also reported for a passage reading task, whereas the present acoustic study measured intensity for a sentence reading task. Tjaden et al. (2013a) reported no difference intensity differences between habitual and slow speech.

Some studies that have included neurologically healthy talkers has reported reductions in intensity in slower speech (Kleinow et al., 2001; Wohlert & Hammen, 2000). Wohlert and Hammen (2000) found that talkers produced similarly reduced speech intensity compared to their habitual levels when speaking at slower rates as when prompted to speak in a softer voice. That is, being instructed to slow down resulted in softer speech.

The DBS group in the present study, who demonstrated similar baseline intensity, showed the greatest decrements in intensity in the slow speech. Speech intensity is a measure that has been shown to improve in some instances following DBS (Tripoliti et al., 2008; Tsuboi et al., 2014), but reasons for increased detriment are not presently known.

Speech intensity has been shown to decline in some dual-task paradigms, particularly those that involve a cognitively effortful task (e.g., tracking movement on a screen) while speaking (Ho et al., 2002). However, other sorts of dual-tasks, such as walking and talking or hand-grip tracking and talking, have been found to be associated with increased speech intensity (Adams, Winnell, & Jog, 2010; McCaig, Adams, Dykstra, & Jog, 2016). McCaig et al. (2016) suggested that tasks such as walking or standing demonstrated an “energizing” effect on conversational speech intensity. It could be the case that focusing on modifying one’s rate of speech to such extremes acts as a cognitively demanding dual task in some sense, more akin to visual tracking. The finding that intensity did not show similar decreases in fast speech, however, does not support this. Anecdotally, many participants commented that slowing their speech down to the extent that they did was very difficult and required substantial concentration.

It could also be the case that slow speech places a greater demand on the respiratory system, and lower speech intensity may be a compensatory mechanism used to maintain continuous respiratory output across an utterance during slow speech. Studies have shown, for example, that speech breathing is affected by utterance length (Huber, 2008; Sperry & Klich, 1992; Winkworth, Davis, Ellis, & Adams, 1994), and in particular for older speakers (Huber, 2008). This possibility is described more below, taking into account the voice quality findings as well.

5.3.2.2 Voice quality

The relationship between speech rate and speaker group was complex for HNR. In effect, the OC and PD groups patterned together, while the DBS group demonstrated marked differences from both. The OC and PD groups sustained the same or better voice quality as their speech rate slowed, and worse voice quality as it quickened. The DBS group, on the other hand, demonstrated the opposite pattern, namely related to poorer voice quality in slow speech. An interaction with vowel height, which reflects the height of the tongue in the oral cavity during vowel production, showed that this pattern was more extreme for the DBS group during the production of high vowels (e.g., /i, u/). That is, they demonstrated a clear pattern of voice quality *decline* in slow speech and *improvement* in fast speech.

Previous accounts of voice quality in PD have reported conflicting findings for signal-to-noise ratio (i.e., HNR, a metric of vocal hoarseness). While some studies have reported that lower (e.g., worse) HNR values in PD (Cushnie-Sparrow et al., 2018; Little, McSharry, Hunter, Spielman, & Ramig, 2009; Oguz et al., 2006; Ramig, Titze, Scherer, & Ringel, 1988; Rusz, Cmejla, Ruzickova, & Ruzicka, 2011; Silva, Gama, Cardoso, Reis, & Bassi, 2012; Tanaka, Nishio, & Niimi, 2011; Yücetürk, Yılmaz, Eğrilmez, & Karaca, 2002), others have reported finding no differences (Bang, Min, Sohn, & Cho, 2013; Gamboa et al., 1997; Graças, Gama, Cardoso, Lopes, & Bassi, 2012; Hertrich & Ackermann, 1995b; Jiménez-Jiménez & Molina, 1997; Midi et al., 2008). The present results are consistent with the latter at the group level, with all other factors held equal. Decreased HNR has also been suggested as an acoustic marker of aging, due possibly to laryngeal and musculature changes in the aging voice, or possibly a side effect of common medications used by the aging population (Ferrand, 2002).

Previous literature has suggested that there may be relative improvements in measures of voice quality such as jitter, shimmer, HNR, and tremor following DBS (D'Alatri et al., 2008; Xie et al., 2011). In a review of speech metrics for the evaluation of speech changes following DBS, Weismer et al. (2012) suggested that measures of voice quality and intensity may show relative improvements, but are often not accompanied by parallel improvements in intelligibility following DBS. While it is not possible to know how the speech and voice characteristics changed following DBS for talkers in the present study, it was observed that this group demonstrated marked overall differences from the other groups. The observation of a decline in slower speech for the DBS group indicates that they may have been producing slow speech with greater laryngeal strain or breathiness, giving rise to more acoustic noise in their voice.

High vowels in general are produced by raising the height of the tongue, in turn eliciting more laryngeal tension (Honda, 1983) and increasing the fundamental frequency of vocal fold vibration (i.e., higher pitch; MacCallum, Zhang, & Jiang, 2011; Fant, 1970; Higgins, Netsell, & Schulte, 1998; L. A. Ramig & Ringel, 1983). Conversely, low vowels are produced at a lower fundamental frequency due to slower vibration of the vocal folds. While vowel-specific voice quality patterns are less clear, evidence suggests that this lower frequency in the production of low vowels introduces more noise into the acoustic signal, accounting for lower HNR compared to high vowels (i.e., poorer voice quality; MacCallum et al., 2011), which is consistent with the findings of the present study.

Laryngeal tension/resistance may be employed to different extents along a continuum, and this in turn could differ across speakers and the specific modifications they make to increase or decrease their speech rate. It could be the case that slight increases or decreases to laryngeal resistance impacted the speaker groups in the present study differently, too. For example, slight increases in resistance may be associated with limited change in voice quality in an unimpaired speaker, but worse voice quality in a speaker with more severe voice impairments. Clustering the participants into groups based on their baseline voice features, rather than on their treatment status, would be one way to better understand these relationships.

As previously mentioned, little literature has systematically explored speech elicited in a *faster* condition in talkers with dysarthria. Some studies have demonstrated that faster than habitual speech is associated with increases in speech naturalness (Dagenais et al., 2006; Logan et al., 2002), but more research is necessary to evaluate the relationship between voice quality and speech naturalness. In a group of 33 Cantonese speakers with dysarthria (including 13 with PD), Whitehill et al. (2004) found no relationship between voice quality acoustics of jitter, shimmer, and SNR, and perceptual measures of speech acceptability. Whitehill et al. (2004) did, however, find that acceptability demonstrated a significant, positive relationship with fundamental frequency. That study did not explore changes in rate, but the relationships among pitch, naturalness, and speech rate suggest that this may be a worthwhile avenue of further investigation.

It could be the case that slowing down one’s speech is associated with multisystem changes in order to achieve the slow rate target. The individuals who demonstrated quieter speech with a higher voice quality in slow speech could have been doing so by decreasing their airflow and increasing laryngeal resistance in order to conserve respiratory airflow over the production of the utterance. These laryngeal changes may be associated with improvements in HNR for the unimpaired or less impaired speakers, but not for the more severe DBS group.

To summarize the acoustic findings of Experiment 1 and address RQ2 (*what are the acoustic changes across groups and speech rates?*), the present study demonstrated that, overall, people with PD and older healthy adults made similar changes to their speech along a continuum of speech rates from very slow to very fast, with some notable differences. Talkers with PD demonstrated abnormal articulatory speech characteristics for voice onset time and vowel centralization but demonstrated similar degrees of phonatory characteristics (intensity and voice quality) as controls. Additionally, speakers with PD and DBS demonstrated poorer articulation but similar voice characteristics compared to the others at baseline, but the modification of speech rate was associated with changes in their voice quality that were not observed in the non-DBS PD or control groups. The present findings also suggest that speech rate led to changes in acoustic distinctiveness for the older healthy controls, but not for the individuals with PD (with or without DBS).

The following sections will address RQ3 (*what are the changes in intelligibility that arise across groups and speech rates?*), RQ4 (*what are the changes in intelligibility that arise across speech rates and speech tasks?*), and RQ5 (*what is the relationship between speech acoustics and intelligibility?*).

5.4 RQ3: Group differences in intelligibility along a speech rate continuum

The results for Experiments 2 and 3 will be presented jointly in order to answer RQ3 (*what are the changes in intelligibility that arise across groups and speech rates?*) and RQ4 (*what are the changes in intelligibility that arise across speech rates and speech tasks?*). As with RQ2 (*what are the acoustic changes across groups and speech rates?*), the most relevant analyses are those that involved group, rate, as well as any additional mediating factors. Differences between tasks will be discussed in a more qualitative manner, as separate analyses were run for the sentence reading and conversational speech samples. All results will be discussed in the context of the literature.

In essence, speech rate was found to have a more consistent, predictable effect on sentence intelligibility estimation than on phoneme identification, supporting the original hypotheses. Slower rates led to higher estimates of intelligibility for the PD groups for sentence and conversational intelligibility. The effects were in the same direction for the control speakers, but less pronounced (i.e., a flatter overall slope). The DBS group demonstrated more variability and fewer significant effects but overall patterned in similar ways. There was a trend for sentence intelligibility to show a slight increase in faster speech for some participants in the DBS group. Phoneme intelligibility, on the other hand, did not demonstrate such clear patterns, and was largely most intelligible at habitual rates for all groups. Slower speech was typically associated with lower phoneme intelligibility compared to faster speech, but this varied by groups and phonemic categories. Overall, the sentence intelligibility results supported the original hypotheses (that slower speech would be rated as more intelligible), while the phoneme identification results did not.

The impact of speech rate on consonant intelligibility, on the whole, did not vary by group, as evidenced by the lack of significant two- and three-way interactions between group and speech rate. Trends in the empirical data support this, as reported in Figure 31; in particular, there are clear drops for all groups in slower speech for both voiced and voiceless stops. Greater change is noticeable for voiced stops, but still the effects are minimal.

Vowel intelligibility patterned in a similar way to consonant intelligibility, such that habitual rates of speech were associated with the highest intelligibility, but differences between the groups emerged across the speech rate continuum.

The three-way interaction between group, rate, and vowel height for the phoneme identification task revealed a complex pattern. In essence, front and back vowels demonstrated opposite patterns for all groups. All groups but the DBS group were, overall, most intelligible at their habitual rates of speech, and the DBS group demonstrated some improvements in the production of back vowels (i.e., /ɑ, u/) at *faster* rates compared to slower rates.

5.4.1.1.1 Asymmetry between sentence and phoneme intelligibility

The results from the present study are largely consistent with what Yorkston et al. (1990) found in a study that explored two slower rates for 8 speakers with dysarthria (half of whom had HkD). The authors of this study found that slow speech, specifically 60% of an individual's habitual speech rate, was associated with gains in sentence intelligibility, but found no improvements in consonant or vowel accuracy in slow speech. Yorkston et al. (1990) proposed that this discrepancy could be due to the fact that the task of understanding a sentence can be facilitated by linguistic and contextual cues. That is, listeners may be more likely to correctly guess a word, even if the sentences are semantically anomalous. In a highly controlled phoneme identification task, only cues from the surrounding phonetic environment are available.

While the present results do, to some degree, replicate the findings of Yorkston et al. (1990), some notable differences between the studies are worth mentioning. Firstly, sentence intelligibility in the present study was an intelligibility estimation task, whereas it was a sentence transcription task scored as percentage of words correct in the Yorkston et al. (1990) study. Therefore, the suggestion that listeners may be able to "guess" words when listening to sentences must be understood slightly differently here. A higher rating may have indicated that the listener believed they understood more of the sentence, but whether or not this corresponded exactly to a concept of words correct remains to be seen. That is, intelligibility may additionally reflect concepts of listener confidence or effort (Maruthy & Raj, 2014; Yorkston et al., 1996b; cf. Hustad, 2007). The finding that the young healthy controls were rated as more intelligible in slow speech was unexpected and may be related to other aspects of speech the listeners were attending to. Recent work, nevertheless, has shown a close relationship between sentence estimation and transcription tasks (Adams, Dykstra, Jenkins, & Jog, 2008; Enos et al., 2018; Stipancic et al., 2016; Tjaden & Wilding, 2011a).

Another important difference is that the stimuli for the phoneme intelligibility task in the present study were mixed with noise at +3 dB; this was not done in Yorkston et al. (1990), but has been reported as a means to minimize ceiling effects in other studies (Ferguson & Kewley-Port, 2002; Kuo et al., 2014; McAuliffe et al., 2009). It should be noted that previous studies have often reported mixing noise at -3 dB SNR. Pilot work prior to the present study suggested that this level was too difficult, and so a positive SNR was chosen instead.

Nevertheless, it is possible that the presence of this noise made the task of phoneme identification too difficult for the listeners. In fact, the average group proportions of phonemes correct for the PD and DBS groups across all rates ranged overall from 41% to 51% for consonant accuracy, and 55% to 68% for vowel accuracy (with averages of 50% and 65% in habitual speech). For younger and older controls, consonant accuracy ranged from 55% to 86% and 67% to 85% for vowels (with habitual speech averages of 79% and 84%, respectively). The goal of the noise was to minimize a ceiling effect, which clearly was achieved, but at the risk of a floor effect, i.e., that the task became *too* difficult. It is important to remember though that this was an open set transcription task, and that a score of "incorrect" gave no weight to whether the consonant shared important phonetic or phonological features with the target (i.e., voicing, place, manner). The finding that these results pattern with those of Yorkston et al. (1990) also suggest that these do not reflect an asymmetry solely attributable to methodological differences.

5.4.1.2 Asymmetry between speech tasks

Yorkston et al. (1990) also reported that the sentence intelligibility ratings were lower than the phoneme intelligibility ratings, suggesting a sort of ceiling effect for phoneme identification results. In the present study, the opposite was observed, most likely due to the presence of noise in the phoneme identification task, though the lack of linguistic and contextual cues in the controlled nonsense word carrier phrases could be implicated. Noise was added, however, because it was assumed that the highly controlled nature of the nonsense word task would lead to ceiling effects in phoneme identification.

While fast speech was generally associated with poorer intelligibility for the sentence and conversational tasks, fast speech was not associated with changes in phoneme identification, in particular for the DBS group. In fact, there were nonsignificant empirical trends for increased intelligibility in fast speech for all three speech tasks for the DBS group (nonsense words, sentence reading, conversation). Exploratory analyses also suggested that some individuals exhibited this trend while others did not. Kuo et al. (2014) reported a similar finding in a study of fast speech (in passage reading) for groups of speakers with PD or multiple sclerosis. On the whole they found that intelligibility tended to decrease in fast speech, but that this was not the case for all speakers, and that some even showed measurable increases in intelligibility at a faster rate.

Van Nuffelen et al. (2009) and Van Nuffelen et al. (2010) found that there were no overall improvements in intelligibility following voluntary rate reduction, as captured by VAS estimation of sentence reading. The results of Experiment 2 are supportive of this finding, but Experiment 3 demonstrated clear trends of increased intelligibility in slower speech (including even the slowest speech) across groups. Notably, these comparisons were significant at both ends of the continuum as well as at the less extreme points. This was captured by the finding that all speech rate comparisons (S3, S2, F2, F3) significantly differed from habitual for the sentence reading intelligibility estimation task.

Exploring the effect of pauses was beyond the scope of this paper but is another important factor of consideration. Previous work has demonstrated that the presence of more frequent or longer pauses alone do not account for gains seen in intelligibility in people with dysarthria talking at slower rates (Hall, 2013; Hammen et al., 1994). The frequency and duration of pauses were not included in the present analysis, but it is possible that more pausing or word boundary separation could have nevertheless contributed to the sentence and conversational intelligibility in some way. Further analysis of the stimuli would be needed to account for this possibility.

5.4.1.2.1 Phonological variables

5.4.1.2.1.1 Vowel backness

Back vowels were transcribed with overall greater accuracy than front vowels in the present study. Front vowels are associated with greater lingual advancement, and back vowels with lingual retraction (Kent et al., 1999) and are distinguished acoustically by the second formant. The ratio between second formant values for the high front vowel /i/ and the high back vowel /u/ are known to be sensitive measures that relate to intelligibility in talkers with HkD (Rusz, Cmejla, et al., 2013; Sapir et al., 2007), indicating abnormalities along this distinction.

There is relatively little literature on vowel-specific vulnerabilities in dysarthria. The current finding of greater transcription accuracy for back vowels is consistent with a recent study that found, for a group of individuals with ALS, the vowel /ɑ/ was identified by listeners with higher accuracy than most other vowels (Lee, Dickey, & Simmons, 2019). For the talkers with severe dysarthria in that study, listeners tended to misidentify the high front lax vowel /ɪ/ as /ɛ/ or /ɑ/.

Tjaden and Sussman (2006) found that listeners were able to identify vowels produced by dysarthric speakers when given auditory information from the preceding consonant They reported a trend for the high back vowel /u/ to be more accurately identified the high front vowel /i/ for the PD speakers (though this was only statistically significant for vowels following /s/). This too is consistent with the present study. In *contrast* to the present findings other studies have reported similar degrees of accuracy for front and back vowels (Ferguson & Kewley-Port, 2002; Lansford & Liss, 2014). Previous literature in healthy talkers has pointed to greater accuracy of *high* vowels /i/ and /u/ (Ryalls & Lieberman, 1982).

Another possible explanation relates to differences in centralization of front versus back vowels. In a series of studies that included speech from 45 speakers with intelligibility of varying etiologies, Lansford and Liss (2014a; 2014b) explored 11 different acoustic vowel metrics to explore their effect on perceptual identification and dysarthria classification. They chose to include metrics of both front and back vowel dispersion, anticipating that there may be distinct patterns of compression when vowel space is reduced. While they found mean, rather than front or back, dispersion to be most related to perceptual accuracy, this raises an interesting possibility for the asymmetry noted in the present study. It could be that front vowels were subject to greater degrees of centralization making them more difficult to identify (Bang et al., 2013), especially in faster speech. Further research is needed to understand vowel-specific changes.

Reduced F2 transitions are a salient acoustic marker of dysarthria (Weismer, 1991; Weismer & Martin, 1992; Weismer et al., 2012). While F2 transitions were not an acoustic variable in the current study, the literature on F2 impairments provide context for the current findings. Impaired F2 slopes indicate abnormalities in lingual advancement as a speaker’s tongue moves forward during the production of a diphthong such as /aɪ/. Articulatory slowness in tongue dorsum movements have also been demonstrated though kinematic evidence in vowel production of individuals with PD (Yunusova, Weismer, Westbury, & Lindstrom, 2008). Taken together, these findings could support the possibility of increased centralization in front vowels in faster speech.

Vowel *height* (i.e., high /i, u/ versus low /æ, ɑ/) has also been found to impact speech intelligibility (Bunton & Weismer, 2001; Lee et al., 2019). Vowel height was not included in the present analysis as a predictor based on exploratory plots that indicated a greater degree of confusability between front/back vowels compared to high/low vowels. However, this is not an observation that can be concluded based on the statistical analysis at this time.

Phonemes have different relative levels of intensity, representing a span of 28 dB between the highest intensity phoneme (mid-back unrounded vowel /ɔ/) to the lowest (voiceless interdental fricative /θ/; Fletcher, 1953; Lawson & Peterson, 2011). In particular, the high front vowel /i/ has the lowest relative power of all the vowels in English (220). Of the vowels included in the present study, /ɑ/ has the highest (600; Fletcher, 1953). Furthermore, low vowels tend to be produced with greater intensity than high vowels. It is possible that this notion of relative power played a role in the listeners' ability to understand, particularly in the presence of added noise.

Two other possibilities related to the present methodology, specifically to prevalence of sounds represented and pronunciation variants, are also considered here. Regarding the former, while the vowels in the subsequent analysis were balanced, the total stimuli set the listeners heard was not. Critically, it contained more instances of the /ɑ/ (low back) vowel. It is possible, then, that listeners picked up on this pattern and were more likely to correctly guess /ɑ/. If this were the case, it would be expected that of the back vowels, /ɑ/ would be transcribed with greater accuracy than /u/ and all other vowels. Closer inspection suggested that this was not a factor in the present analysis; /ɑ/ and /u/ were correctly transcribed with 83% and 81% of cases, respectively, while the front vowels /i/ and /æ/ with 57% and 58% (these numbers correspond to the positive predicted value). Further statistical testing would be needed to rule this possibility out entirely.

Regarding pronunciation variants, there are two sources of potential confounds: the orthography of the task, and dialectal characteristics. The novel words containing the vowel /æ/ were orthographically written as "aCad." Anecdotally, it was noted that some speakers were inclined to produce /æ/ as /ɑ/. Speakers read aloud the list of words first and were informed on the intended pronunciation. However, during the task, sometimes speakers would produce other variants. In the case that these were clearly related to orthography, the researcher would have them repeat the trial again, reminding them of the intended target. In cases where it was not a clear misreading, the researcher would not interrupt them. In Canadian English, the vowel /æ/ tends to be produced in a lower, more central position due to a process known as the Canadian Shift (Clarke, Elms, & Youssef, 1995). This is more prevalent in younger speakers and also is known to affect short front vowels /ɪ/ and /ɛ/ (Boberg, 2005; Labov, Ash, & Boberg, 2008), which were not included in the present analysis. It is possible then that speakers were producing more "backed" front vowels, causing them to be more confusable with phonologically back vowels. In novel words this may have led to greater confusion for the listeners. Given that there were significant interactions between vowel backness, speech rate, and group suggests that even if this were the case, factors related to dysarthria, aging, and speech rate are all to some degree also at play.

5.4.1.2.1.2 Stop consonant voicing

Across the groups, voiceless stops were transcribed with greater accuracy than voiced stops. The hierarchy of accuracy (YC > OC > PD > DBS) was maintained for this contrast, and younger controls demonstrated less of a difference in the accuracy of voiced and voiceless stops. This is seen as greater overlap for the controls in Figure 31.

While the finding of increased voicing into closure in older adults and individuals with PD (Weismer, 1984a) would seem to support a greater likelihood of *voiceless stops* being confused for *voiced* stops, this was not the case. Consistent with the present findings, though, other researchers have also reported a high incidence of devoicing of voiceless stops (Antolík & Fougeron, 2013; Bunton & Weismer, 2002). In particular, phonetic intelligibility testing of 35 speakers, including 10 with PD, demonstrated that the voiced-voiceless distinction was among the most frequent errors for all groups with dysarthria, as well as healthy older females (Bunton & Weismer, 2002). The authors of this study found that 78% of voicing errors were attributed to listeners incorrectly perceiving voiced consonants as voiceless, even in the case where voicing during closure of voiceless stops was identified. Similar error patterns were found in the healthy geriatric group as well as in talkers with dysarthria. While a full phonetic profile was not completed in the present study, the current findings corroborate this pattern, which Bunton and Weismer (2002) hypothesized may be related to increased laryngeal musculature stiffness due to aging, rather than a process specific to dysarthria.

5.4.1.2.2 Sentence length

The finding of higher intelligibility for shorter sentences is consistent with previous literature that demonstrates a detrimental effect of increased sentence length on intelligibility in dysarthria (Allison, Yunusova, & Green, 2019; Beverly et al., 2010; Hustad, 2007; Yunusova et al., 2005) and healthy adults (Bradlow et al., 1995; Huber, 2008). There is likely a trade-off between the linguistic and contextual content afforded by longer utterances over single words or short phrases (Carter, Yorkston, Strand, & Hammen, 1996; Hammen, Yorkston, & Dowden, 1991; Hustad, 2007), and the increased motoric demands of longer utterances, such as increased respiratory support and articulatory demands (Allison et al., 2019; Huber, 2008).

One curious discrepancy in the present data was that while, for the most part, there was a predictable linear downward trend regarding intelligibility and sentence length, the 8-word sentences were transcribed with higher intelligibility than some of the shorter sentences. This was more or less consistent across groups, but more pronounced for the PD and DBS groups, as can be seen in Figure 34 A. There is visibly less of a decline in fast speech for the 8-word sentence, as seen in Figure 34 B. Reasoning for why this might be is purely speculative at this point, but perhaps it is worth considering that there may be an “ideal” sentence length at which top-down linguistic information is maximally facilitative and sentence complexity is minimized.

Listeners rated the conversational speech samples, which were often observed to include longer phrases (though an analysis of sentence content is beyond the scope of this thesis) similarly to the sentences. This was evidenced by the lack of statistically significant differences between the intelligibility of the two tasks across all speech rates. While this is inconsistent with several reports of greater speech impairment in PD being more noticeable in more spontaneous speech (Bunton & Keintz, 2008; Ho et al., 1999; Sidtis et al., 2012, 2010), it is consistent with at least two studies that have demonstrated this may not always be the case (Bunton & Keintz, 2008; Tjaden & Wilding, 2011a). Bunton (2008) found that talkers with PD were rated more poorly in spontaneous speech compared to read speech, but that this task difference disappeared when the participants were asked to perform a motor task simultaneously (screwing a nut on a bolt). As previously stated, it is possible that attending to one's speech rate could be considered a kind of dual-task as well, which would perhaps level the differences between conversational and sentence reading differences that would impact intelligibility.

Even if the conversational speech samples did demonstrate greater objective impairments (which is not known from the present analyses), the finding that intelligibility ratings were similar to those during the reading task could be reflective of the benefit on narrative content on intelligibility. Previous research suggests that narrative context may facilitate listener intelligibility by providing additional contextual information (Drager & Reichle, 2001; Hustad, 2007).

While the sentences and conversational samples were not directly compared across the rate continuum, visual comparison of Figure 35 and Figure 36 suggests that the DBS group appeared to be rated as less intelligible in the conversational samples, as well as to demonstrate greater improvements in slower conversational speech. More severe dysarthria, as is the case in the overall less-intelligible DBS speaker group in the present study, may not be subject to the same gains of narrative speech (Hustad & Beukelman, 2001). It is also possible that there were differences in the richness of linguistic and contextual information impacted by discourse and cognitive factors (the reader is directed to the review by Altmann & Troche, 2011). Spontaneous speech production has also been shown to be more sensitive to speech impairments compared to more formal speech tasks in PD in general (Bunton & Keintz, 2008; Ho et al., 1999; Sidtis et al., 2010) as well as in talkers with DBS (Sidtis et al., 2012).

5.5 RQ5: Relationship between speech acoustics and intelligibility along a speech rate continuum

The final research question pertained to the relationship between the acoustic variables studied in Experiment 1 and the intelligibility results found in Experiment 2 (the nonsense word transcription study). Speech intelligibility, measured at the word level, was modelled as a function of these five acoustic variables, as well as group, rate, speaker sex, and consonant voicing. Results indicated that, of these five acoustic variables, all of which demonstrated sensitivity to speech rate, all but one were significantly associated with intelligibility.

Specifically, VOT (with stop voicing taken into account), vowel centralization, speech intensity, and voice quality were significantly associated with word-level intelligibility. Voicing during closure was not.

In the present study, both measures of voice and articulation were found to be significantly associated with intelligibility. These findings are consistent with studies that have explored multiple speech system domains and acoustic variables (Kim et al., 2011; Whitehill et al., 2004; Yunusova et al., 2005). Some reports, though, have found suprasegmental variables to be more important contributors to intelligibility (Kim et al., 2011) and acceptability (Whitehill et al., 2004) in dysarthria.

The purpose of this section was to account for *explanatory* correlates of speech intelligibility in a highly controlled task when speaker group and speech rate were taken into account. It is not meant to be an exhaustive exploration of these factors, on the contrary, there are several avenues of further investigation that are warranted given the results. Each of the acoustic variables in this study will be considered in the context of the literature on acoustic predictors of speech intelligibility in PD, as well as in light of the findings of Experiments 1 and 2 in the present study. Implications for the findings in Experiment 3 will be discussed.

5.5.1.1 Voice onset time

VOT on its own was not a significant predictor of intelligibility, but its interaction with stop consonant voicing was. Previous literature on VOT's role in intelligibility in dysarthria is sparse and inconsistent. Bunton & Weismer (2002) found that a large proportion of phonetic error patterns in individuals with and without dysarthria (including individuals with PD) were related to consonant voicing contrasts, but considerable overlap in VOT suggested that VOT could not explain the perceptual errors. This was consistent with other stop consonant voicing cues included in the Bunton & Weismer (2002), namely, F1 onset, voicing during stop closure, and preceding vowel duration, which also showed considerable overlap and did not significantly differ across correctly versus incorrectly identified tokens.

Another study demonstrated that VOT was a significant predictor of intelligibility for young Mandarin-speaking adults with dysarthria secondary to cerebral palsy (Liu, Tseng, & Tsao, 2000), but comparison with this study is difficult due to differential linguistic, age-related, and dysarthria-specific factors. One previous study found that the *voiceless interval duration*, an alternative metric to VOT that takes into account the stop closure and VOT for voiceless stops, *was* a significant predictor of intelligibility for a group of dysarthric speakers that included VOT (Kim et al., 2011). The finding of the current study that VOT by itself was not a predictor, but the interaction between VOT and voicing was, suggests that the relationship between VOT and intelligibility is highly modulated by phonological contrasts. VOT is a robust stop voicing cue (Lisker & Abramson, 1964), and explaining perceptual errors related to a voicing contrast must take this into account.

In the present study, VOT was also closely related to speech rate, demonstrating a fairly linear trend with largest VOTs (for voiced and voiceless stops) in slow speech, and shortest VOTs in fast speech; this was unsurprising. VOT was also found to be *longer* for the clinical groups (in particular for the DBS group), and this too was consistent across voicing categories. An asymmetry across the OC and clinical groups existed in the group by rate interaction that showed that the voicing contrast for the OC group was modulated by speech rate, whereas the clinical groups demonstrated a fairly consistent contrast in slow and fast speech. It was suggested that this may be largely in part due to changes in voiced VOT present in the clinical groups across speech rates, a finding that is uncommon for healthy talkers (Kessinger & Blumstein, 1997; Miller et al., 1986, 1997; Summerfield, 1981).

Relating back to intelligibility, Panel A of Figures 37 and 38 demonstrate that as voiced VOT increases (i.e., becomes more voiceless-like), intelligibility decreases for voiced stops. This trend is particularly evident for the PD groups, and less so for the OC group.

5.5.1.2 Voicing during closure

Neither voicing during closure nor its interaction with voicing, were significant predictors of intelligibility in the present study, but this null effect warrants discussion. It is also important to recall that these utterances were mixed with background noise, which likely obscured the perceptual salience of VDC. VDC was found to consistently vary across speech rates, with a greater degree of voicing throughout the entire closure occurring in faster speech for all groups. VDC was also found to occur more for the PD and DBS groups, and in voiced compared to voiceless stops. VDC has been reported to be related to voiced stop identification in healthy talkers (Bradlow et al., 1996), but does not appear to be related in dysarthria (Bunton & Weismer, 2002). A small body of research on acoustic predictors of intelligibility of foreign-accented speech has also found a relationship between VDC and perception of voiced stop consonants in Mandarin (Hayes-Harb, Smith, Bent, & Bradlow, 2008; Xie & Fowler, 2013), but not for all talkers or listeners.

It could be the case that VDC, while a sensitive metric to rate and group, was not as salient a cue for older talkers in general. Comparisons to the younger speakers, who would be hypothesized to demonstrate less VDC overall (Weismer, 1984a) and who were found to be more intelligible would shed more light on these processes. Bunton & Weismer (2002), who found that VDC and other acoustic stop voicing cues did not reliably map on to perceptual error patterns, suggested that listeners may develop a "tolerance" (pp. 236) for acoustic information that age-related, such as increased VDC, or weight it less heavily as a perceptual cue.

5.5.1.3 Vowel articulation

Vowel centralization, as captured by QVAI, was found to be a significant predictor of intelligibility. This is consistent with several studies that have found acoustic vowel metrics to be related to intelligibility in dysarthria (Feenaughty et al., 2014; Kim et al., 2011, 2009; Lansford & Liss, 2014; McRae et al., 2002; Tjaden et al., 2013a; Tjaden & Wilding, 2004; Yunusova et al., 2005). In particular, measures that encapsulate vowel overlap may be more predictive than those that measure vowel *space* (H. Kim et al., 2011; Neel, 2008). The influence of QVAI in the present study lends evidence to the predictive value of vowel overlap, as it reflects acoustic space occupied by the four corner vowels (Karlsson & van Doorn, 2012; Sapir et al., 2010).

QVAI was also found to vary closely with speech rate and speaker group in Experiment 1. Specifically, faster speech was associated with more vowel centralization. While there was a trend for larger QVAI in slower speech, this was only marginally significant at the slowest rate. The results of Experiment 1 demonstrated that QVAI became much more variable at slow rates, as seen in Figure 21 B. QVAI was largest for the healthy controls, followed by the PD group, and smallest for the DBS group. The relationship between QVAI and intelligibility is probably most easily explainable for the OC and PD groups. Both of these groups demonstrated more vowel centralization in fast speech and decreased intelligibility in fast speech. The linear relationship between QVAI and intelligibility makes sense in this case.

On the other hand, the DBS group also demonstrated more vowel centralization in fast speech but did not become less intelligible with regard to the vowel accuracy measure. In fact, there was a trend for the DBS group to become more intelligible in fast speech.

5.5.1.4 Vowel intensity

Vowel intensity was positively associated with speech intelligibility. It is important to reiterate that the audio clips presented to the listeners were scaled to 70dB to remediate differences related to volume and were presented in babble noise at +3SNR. This scaling procedure was carried out on the intensity of the whole carrier phrase.

The acoustic analyses, however, were carried out on the unaltered files (unscaled and without noise), and intensity was calibrated using the technique described in Section 3.2.3. Therefore, the finding that intensity was significantly associated with increased intelligibility does *not* indicate that this was related to the intensity of the speech signal the listeners heard, or the relative volume of the speech signal to the noise.

While the scaling procedure was attempted to account for large fluctuations and differences in loudness and to be able to better assess segmental contributors of intelligibility, *relative* intensity across the phrase would have largely remained intact. Inherent relative intensities across vowels, as described above in Section 5.3.2.1, could have facilitated listeners' intelligibility of vowels with greater relative intensity, such as /ɑ/. The effect of relative intensity on the nearby consonants may also have had an effect.

It is worth acknowledging that fast speech compared to *habitual* speech was not associated with increased intensity, but fast compared to *slow* speech was (this is indicated by the values reported for the post-hoc pairwise comparisons in Figure 25.

Loud speech is a common therapeutic target in PD (see reviews in Yorkston et al., 2007; Atkinson-Clement, Sadat, & Pinto, 2015), given the prevalence of hypophonia (Adams & Dykstra, 2009), and has been found to be associated with intelligibility gains (e.g., Tjaden & Wilding, 2004; McAuliffe et al., 2017; Tjaden et al., 2013b, 2014b; Yorkston et al., 2007).

Few studies, though, have explored the precise nature of speech intensity on intelligibility. Adams et al. (2008) found there to be a close relationship to conversational speech intelligibility and speech-to-noise levels in both healthy talkers and those with PD and hypophonia; that is, the louder the talkers were above background noise levels, the better they were understood. Kim et al. (2011) found that speech intensity *range* was not predictive of intelligibility in dysarthria (including HkD), but they did not look at speech intensity as a static measure.

Altering the loudness of stimuli for the listener (e.g., synthetically amplified speech rather than speech produced at a greater intensity) is not necessarily associated with increases in intelligibility with dysarthria (Kim & Kuo, 2012; Turner, Martin, & de Jonge, 2008; cf. Iddon, Read, & Miller, 2015). Intelligibility ratings of speech produced at a loud volume versus amplified versions of the same speech have shown discrepancies, leading some researchers to conclude that loudness alone must not be responsible for the observed intelligibility gains (Neel, 2009). The use of speech amplification devices as a form of augmentative intervention, however, have demonstrated that some individuals are more intelligible when their acoustic signal is amplified to the listener (Iddon et al., 2015). While the current findings do not map on to these paradigms directly, the overall conclusion is that speech produced at higher intensities was indeed associated with increased intelligibility in the present study, but it is unclear to what degree that is associated with loudness itself, versus other speech changes that accompany louder speech production.

The literature generally shows that loud speech, like slow speech, is associated with increased acoustic vowel space (McAuliffe et al., 2017; Tjaden et al., 2013a; Tjaden & Wilding, 2004), acoustic consonant working space (Tjaden et al., 2013a, 2013b, 2014a; Tjaden & Martel-Sauvageau, 2017; Tjaden & Wilding, 2004), and articulatory space (Darling & Huber, 2011). In the present study, slow speech was associated with quieter speech (compared to habitual speech) and fast speech was associated with louder speech (compared to slower speech).

Kuo et al. (2014) reported that talkers with PD or MS increased their speech intensity at a faster rate, and that the magnitude of change was greater for those talkers who also demonstrated increased intelligibility at a faster rate as well (7 dB versus 2 dB for those who were not more intelligible). Louder speech intensity at faster speech rates has also been reported in young healthy talkers (Dromey & Ramig, 1998; Wohlert & Hammen, 2000) as well as in one study of an individual with dysarthria secondary to traumatic brain injury (D'Innocenzo et al., 2006).

It is possible that an increased speech rate was, for at least some individuals, associated with increased speech motor effort, such as changes to phonatory, respiratory, or articulatory effort. Increased phonatory effort is a key component of the Lee Silverman Voice Treatment (LSVT; Ramig, Fox, & Sapir, 2004), which aims to bring about multisystem speech changes by training individuals with PD and HkD to speak louder. Several studies have documented evidence of improvements in respiratory, laryngeal, and articulatory parameters (Dromey, Ramig, & Johnson, 1995; Dumer et al., 2014; El Sharkawi, 2002; Ramig & Dromey, 1996; Sapir et al., 2010), including in some individuals with DBS (Spielman et al., 2011). Cannito et al. (2012) found, however, that two of eight speakers did not show improvements in speech intelligibility following LSVT, even though one of the two did demonstrate expected increases in speech intensity.

In a study comparing habitual, fast, slow, loud, and quiet speech in people with PD as well as younger and older controls, Kleinow et al. (2001) found that loud speech was associated with the least amount of spatio-temporal variability. Slow speech was associated with the greatest degree of variability for all groups. It could be the case that in the present study, the variability in the quieter speech produced in the slow conditions may have negatively impacted intelligibility. While not a primary outcome measure of their study, Kleinow et al. (2001) also reported that slow speech was associated with a decrease of speech intensity of approximately 2 dB. The older control group *increased* their speech intensity in fast speech by approximately 2 dB as well, but the younger controls and PD group saw much smaller increases, on the order of less than 0.5 dB. These patterns are also consistent with the current study.

5.5.1.5 Harmonics-to-noise ratio

Overall, voice quality, as measured by HNR, demonstrated a significant negative relationship with intelligibility, suggesting that higher transcription accuracy was achieved when HNR was lower compared to higher. This was an unexpected finding, as it was hypothesized that higher signal-to-noise ratios would facilitate a listener's ability to correctly perceive the spoken utterance. This finding does, however, make sense when related back to the inconsistent findings of HNR across speech groups and speech rates. Two of the three groups, the OC and PDs, demonstrated higher HNR in slow speech, but slow speech was also associated with lower intelligibility for these groups. Conversely, the DBS groups demonstrated lower HNR in slow speech, but did not demonstrate a change in intelligibility. There could thus be a cancelling out effect which in essence might point to a three-way interaction that was not accounted for in this analysis. Therefore, while HNR could be positively related to speech intelligibility in habitual speech, the vast range of speech and voice alterations across the rate continuum indicate that, overall, HNR and speech intelligibility were inversely related in this study. The effects of HNR on phoneme-specific intelligibility were not explored (i.e., the results of RQ5 take into account the combination of vowel and consonant intelligibility at the word level).

In a study assessing viable acoustic measures to capture speech changes induced by DBS, Weismer et al. (2012) suggested that voice quality measures (including HNR) may not be ideal candidates. The authors attributed this to the finding that improvements in voice quality have been reported in individuals with PD following DBS, but that these changes are rarely associated with simultaneous improvements in intelligibility (D'Alatri et al., 2008). The results of the present study support this observation.

It is also important to reiterate that these analyses modelled the linear trend. A nonlinear analysis was explored but ultimately abandoned as the nonlinear terms did not significantly improve the models. Nevertheless, a more in-depth investigation of the directionality of relationships between these acoustic variables and intelligibility *within* groups, rates, and individual speakers, would shed more light on the processes that underlie these intelligibility changes.

5.6 Limitations

Several limitations to the present study warrant caution in interpreting the findings. Those not mentioned in the Discussion are presented here. These can largely be broken into the following categories, each of which will be discussed in turn: data collection limitations and methodological/analysis decisions.

5.6.1 Data collection

5.6.1.1 Participant sample

The sample size, particularly for the DBS group, was relatively small. Considering the variability in this clinical population, this should be noted as a limitation to the generalizability of the sample. Furthermore, the distribution of male and female speakers across the groups was uneven, with only four of 22 in the PD group and two of 12 in the DBS group. This is also problematic. Males are diagnosed with PD approximately 1.5 times more frequently than females (Fahn, 2003; Wirdefeldt et al., 2011), but the distribution of females in the present study did not reflect that proportion. Furthermore, there was an even split of male and female speakers in the control groups. As stated throughout the thesis, effects and interactions involving sex (and particularly involving sex and group) should be considered with extreme caution for this reason.

A consideration that must be taken into account when studying older adults (with or without neurological disorders) is other age-related factors that may impact speech analyses in unaccounted for ways. For example, a number of individuals in the present study reported wearing or planning to be fit for dentures. Liss (1990) conducted a study on the speech of very old males (>87 years old), all of whom had dentures. She acknowledged that dentures could affect oral sensory patterns or impact speech production in other ways but speculated that this did not alter their findings because their articulatory acoustic results patterned with previous literature. Here, too, the overall results of fine-grained temporal measures such as VOT patterned with previous work, suggesting that dentures did not necessarily play a role. Further work that investigates other measures of speech production, for example spectral measures, should account for potential differences related to denture use.

Two of the older healthy control participants (OC06 and 0C10) presented with mild perceptible speech abnormalities. One of these (OC06) did wear dentures but reported he did not notice any differences in his speech when he didn’t wear them. The other (OC10) did not wear dentures but did report he had noticed changes in his speech a few years prior, though didn’t attribute this to any diagnosis or event. While these speech changes could be indicative of an underlying, undiagnosed neural abnormality, it is impossible to say with certainty. It was decided to keep these individuals as this may very well be representative of the variability in aging speakers as well.

5.6.1.2 Speech task

Related to the speech task itself, one limitation was its length. Participants were encouraged to take breaks as needed, but the protocol took approximately two to three hours in total. This raises a risk of speech changes that arise not from neurological or age-related motor patterns but from fatigue. With this in mind, the experiment was designed to be counter-balanced to control for order-effects and fatigue.

The speech tasks were conducted in a highly controlled setting so as to limit sources of error and unwanted variability, but it is well known that speech elicited in such highly controlled settings differs from speech in more naturalistic environments (Byrd, 1994; Picheny, Durlach, & Braida, 1989; Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988). In particular, the novelty of the items in the nonsense word task could further have elicited a less natural speaking style in participants. Further work comparing these highly controlled and less controlled speaking environments is needed to better understand the limitations to the general field of disordered speech research.

Relatedly, it is well known that people with PD are particularly affected by external cues (Ho et al., 1999; Weir-Mayta et al., 2017), which a lab setting and novel speech tasks could be considered to be. It was observed, for example, that some of the participants with softer, imprecise speech adopted a louder, clearer voice without prompting during the more formal tasks. This is a variable that is very difficult to account for in this population but is further reason for a need to continue finding ways to incorporate more naturalistic speech tasks in studies.

5.6.2 Methodological decisions

5.6.2.1 Intelligibility tasks

One potential limitation is the impact that the noise had on the listeners' abilities to identify the consonants and vowels in the present study. The choice to include noise was in keeping with previous literature in an attempt to reduce a ceiling effect (Ferguson & Kewley-Port, 2002; Kuo et al., 2014; Maniwa et al., 2008; McAuliffe et al., 2009), but it is possible that the introduction of noise introduces dilemmas of its own. A recent study reported while noise of course has a detrimental effect on intelligibility, the degree of detriment is similar for healthy and disordered speaker groups (Yoho & Borrie, 2018). More work is needed to determine the effect of added noise specifically on segmental identification.

Acknowledging that, the proportion of phonemes correct is similar to that reported in previous studies. For example, vowel accuracy for the PD group at their habitual speech rate was found to be 73% in the present study, and word accuracy was found to be 47%. Lansford and Liss (2014) reported on vowel identification from a sentence transcription task (not presented in noise) and found vowel accuracy for their PD group to be 80%, and 57% for words.

The listeners in this study spent several hours (~10) doing the listening tasks over the course of multiple sessions and weeks. It is likely that there was a learning effect over this time. Perceptual learning is an acknowledged drawback of perceptual studies such as these (McAuliffe et al., 2017). To try to account for this as best as possible, all listeners underwent a brief practice period, and all stimuli presentations were completely randomized for each listener. Inter-rater reliability for the listening tasks was, for the most part, found to be at acceptable levels (though lower than ideal for inter-rater vowel identification in some listener pairs). In the analysis, two approaches were taken to account for listener variability, as described in the methods. In the transcription task, listener was included as a covariate, and item was included as a random effect. In the visual analogue scale estimations, responses were averaged across listeners.

Another potential limitation is what the listeners were attending to during the intelligibility estimation task. While they were told to rate "how understandable" the speech was, they could have been attending to other aspects of the speech signal. This is a common criticism of "subjective" metrics of intelligibility, such that each listener not only has their own internal standard, but also may be attending to different aspects of the signal that are difficult to control for (Miller, 2013; Weismer & Martin, 1992). As mentioned earlier, a growing body of literature suggests a tight relationship between scalar estimates of intelligibility and transcription (Adams et al., 2008; Enos et al., 2018; Stipancic et al., 2016; Tjaden & Wilding, 2011a).

5.6.2.2 Speech production task

Modulating one's rate of speech is a difficult thing to do and, in a clinical setting, often requires extensive training for rate adjustments to be able to generalize to functional communication (Blanchet & Snyder, 2010; Yorkston et al., 1988). The speakers in the present study did not undergo any such training, though they were given the opportunity to practice each rate until they were ready to proceed. The decision to play a recording of their own rate back to them every five to ten trials to help them stay on target was made specifically to make this task easier and increase its internal validity. The decision to look at actual (proportional) rates of speech rather than rate conditions was also made to account for this. Nevertheless, whether the changes observed here would persist after a speaker had more time to *learn* to speak at different target rates is presently unknown.

Speakers were asked to say the same nonsense word items several times over the course of the elicitation. It is well known that words that are repeated demonstrate reduced prominence (e.g., Aylett & Turk, 2004; Bard et al., 2000; Bard & Aylett, 1999; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Fowler, 1988; Lam & Watson, 2010). To account for this, all speakers read aloud the novel words prior to beginning the task. Nevertheless, there still could be an effect of repetition over time that was not taken into account.

5.6.3 Technical drawbacks of the recording procedure

As mentioned in the methods, audio clipping was observed during some speech. This was due in part to individuals' own speech adjustments across the task, but also due to an unforeseen technical difficulty in the calibration procedures. Excessive clipping across a whole utterance, which occurred rarely, was discarded, but most items with clipping were kept for the listener playlists to avoid the intrusion of systematic bias. Including clipping as a covariate in the acoustic model of intelligibility was done as a way to account for this as well. Nevertheless, this is a problem that must be monitored in the future for these data.

Another technical challenge related to the observation of microphone pops in some utterances. In few cases, these were very large, which could have had an effect on the end result of the scaling and noise mixing procedure. That is, utterances with very loud mic pops may have ultimately been presented with lower SNR at the word of interest, because the overall intensity of the utterance would have been affected by the amplitude burst during the microphone pop. These utterances were extremely few (< 0.5% of the data), and so if this impacted intelligibility, its overall contribution would be negligible.

5.6.4 Analysis limitations

The primary models used for the analyses took into account only a subset of the possible variables that could account for the differences seen. For example, speaker group was included, but not any metric of severity (speech or motor). Furthermore, apart from the younger versus older distinction, age was not explicitly analyzed. Other variables, such as cognitive function or depression scores, both of which are known to be prevalent in PD (Reijnders et al., 2008) were not accounted for in the analyses. The modelling decisions were made in order to identify the effects of the primary variables of interest while taking into account the most viable contributing variables from the literature (e.g., for the consonant acoustic metrics, choosing to include the phonological variables of voicing and manner). Nevertheless, there are other variables that likely could have impacted the results further that are not included.

Along these lines, the final analysis, which modelled the effects of five acoustic variables on intelligibility, of course reflects an extremely small subset of all possible acoustic variables impacting intelligibility. Intelligibility is a multifactorial phenomenon, and large-scale sophisticated analyses may be better equipped to identify a more comprehensive set of acoustic variables (e.g., machine learning, principal component analysis, etc.). With this in mind, this particular analysis was conducted on a highly controlled speech task for this very reason; to impose a limit on the number of the most important acoustic variables. Exploring the speech intelligibility of the sentences and conversation are of great interest to this body of work but would necessitate looking at a wider host of linguistic variables (i.e., prosodic, semantic, syntactic, etc.). The highly controlled nature of the task limits the likelihood of influence from these other domains.

5.7 Clinical implications

The current study was derived from the observation that slower speech is a commonly employed clinical goal for individuals with PD and articulatory speech impairments but has shown to have variable and sometimes detrimental effects on intelligibility. Overall, this study found that poorer phoneme accuracy (identified from a carrier phrase) was observed in slow speech, and little to no change in fast speech (in fact, a trend for improvements in the DBS group). Acoustic variables mediating these changes were found to relate to stop consonant and vowel distinctiveness. Speech intensity and voice quality were also found to be related to phoneme accuracy, though in a less straightforward manner. On the other hand, more naturalistic speech was associated with higher speech intelligibility at slow rates, and lower intelligibility at fast rates for all groups.

The DBS group, who was found to have the most severe articulatory and intelligibility impairments, overall patterned differently from the control speakers as well as the non-DBS PD group, including trends for positive speech changes in *fast* speech. Whether these changes are associated specifically with DBS or overall speech severity is presently unknown from the current findings. The variability points to future areas of research investigating different speaker groupings and within-speaker approaches. This is especially relevant for people with PD who are candidates for or have received DBS. As previously stated, further speech impairments following DBS are common, and a better understanding of the generalizability of typical behavioural speech interventions for PD applied to this subgroup is ever more important. There are no current guidelines for speech intervention best practices for these individuals, and there is relatively little research exploring the differences in their responses to intervention, given the well documented variability in speech symptom presentation. More research is needed to fill this gap.

Findings from this study demonstrate a need to be vigilant of speech and voice changes that accompany slower-than-habitual rates of speech, if a slower rate of speech makes clinical sense as a therapeutic goal for a given individual with PD. This study also adds further evidence to the small body of literature that has suggested exploring the use of fast speech as a mechanism for inducing positive vocal changes, even if fast speech is not necessarily a therapeutic goal. Furthermore, it emphasizes the well-acknowledged finding of large interspeaker variability in this population, and points towards a need to explore individual variation in order to better determine candidacy for these behavioural modifications.

5.8 Future directions

A reasonable next step for this work is to further explore individual responses to determine those who fit the group patterns and those who did not. Clustering subgroups, for example by specific speech symptoms (Fletcher et al., 2017b; Tsuboi et al., 2014) or by perceptual scores (Cushnie-Sparrow et al., 2018; Im et al., 2018) would be a promising avenue for this, as would within-speaker approaches (Feenaughty et al., 2014; Yunusova et al., 2005). An important clinical implication from this would be to better tune assessments to determine candidacy selection for rate adjustments (Yorkston et al., 2007).

An investigation of the acoustic patterns across the rate continuum in the spontaneous speech is another area that calls for more research. In particular, determining whether the same set of acoustic variables are associated with intelligibility in more naturalistic tasks would be of great value, adding to the vast body of literature that has attempted to better understand these connections. Further investigation of a larger set of articulatory-acoustic variables across these different speech tasks and rates is also of interest.

Yorkston and colleagues have long acknowledged that there is a close relationship and trade-off between speech intelligibility, speech rate, and speech naturalness (Yorkston et al., 1988, 1999, 1990). An obvious next step would be to collect other perceptual features of the speech elicited across a wide rate continuum to determine how concepts such as naturalness and acceptability map on to the acoustics and intelligibility results. Furthermore, investigating other listener-related factors such as listener effort, both perceptually and physiologically, is of interest. Other acoustic variables related to prosody, such as intonation contours and perceived stress, would also provide important information about these patterns more globally.

A curious finding that adds to a small but growing body of literature (Dagenais et al., 2006; D'Innocenzo et al., 2006; Kuo et al., 2014; Logan et al., 2002) was that fast speech in some cases was associated with improvements in speech for the talkers with dysarthria. More work is needed to determine whether there are particular speaker characteristics that benefit more than others, and whether there is the potential to take advantage of such effects in therapeutic way.

5.9 Conclusions

This series of experiments adds greater detail to the body of literature investigating speech rate modifications in dysarthria. Overall, the results demonstrated that a slow rate of speech, which is often a therapeutic target for talkers with PD and articulatory impairments, was associated with improvements in intelligibility in certain cases. Slow speech was also associated with greater variability, especially at the extreme end of the continuum, and the benefits of slower speech were not evident across all tasks or speakers. Very slow speech was also associated with decreases in phoneme identification. Fast speech was more consistently associated with decreases in intelligibility and acoustic distinctiveness in the expected direction, but also was associated with subtle improvements in certain aspects of voice and speech production. This work adds to the field's knowledge of the large amounts of variability and the complex relationships between speech rate, speech intelligibility, and speech acoustics.

Appendix

Appendix A: Western University Health Sciences Research Ethics Board approval and extension

Western Research

Date: 22 December 2017

To: Dr. Scott Adams

Project ID: 110455

Study Title: The effect of speech rate manipulations on acoustic distinctiveness and speech intelligibility in dysarthria

Application Type: HSREB Initial Application

Review Type: Delegated

Meeting Date / Full Board Reporting Date: 09/Jan/2018

Date Approval Issued: 22/Dec/2017

REB Approval Expiry Date: 22/Dec/2018

Dear Dr. Scott Adams

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above mentioned study, as of the HSREB Initial Approval Date noted above. This research study is to be conducted by the investigator noted above. All other required institutional approvals must also be obtained prior to the conduct of the study.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Intake_Form	Paper Survey	15/Sep/2017	
LOI_AD_2017-12-15_clean	Written Consent/Assent	15/Dec/2017	3
LOI_HA_2017-12-15_clean	Written Consent/Assent	15/Dec/2017	3
LOI_Listeners_2017-12-15_clean	Written Consent/Assent	15/Dec/2017	1
LOI_PD_2017-12-15_clean	Written Consent/Assent	15/Dec/2017	3
MDS-UPDRS_English_FINAL	Other Data Collection Instruments	01/Dec/2017	1
MoCA-Test-English_7_1	Other Data Collection Instruments	01/Dec/2017	1
Recruitment_CCAA-flyer_2017-11-30	Recruitment Materials	01/Dec/2017	1
Recruitment_Classroom_Script_2017-11-30	Recruitment Materials	01/Dec/2017	1
Recruitment_Email_Reminder_Script_2017-12-15_clean	Email Script	15/Dec/2017	1
Recruitment_Email_Script_2017-12-15_clean	Email Script	15/Dec/2017	2
Recruitment_Telephone_Script_clean_2017-11-30	Telephone Script	01/Dec/2017	2
Speech intelligibility estimation	Paper Survey	08/Oct/2017	
Speech intelligibility transcription	Paper Survey	08/Oct/2017	
Speech_Stimuli_Examples	Other Data Collection Instruments	15/Sep/2017	

Documents Acknowledged:

Document Name	Document Type	Document Date	Document Version
Knowles_Prospectus_2017-11-30_revised_clean	Protocol	01/Dec/2017	2

No deviations from, or changes to, the protocol or WREM application should be initiated without prior written approval of an appropriate amendment from Western

Page 1 of 2

Date: 25 November 2018

To: Dr. Scott Adams

Project ID: 110455

Study Title: The effect of speech rate manipulations on acoustic distinctiveness and speech intelligibility in dysarthria

Application Type: Continuing Ethics Review (CER) Form

Review Type: Delegated

REB Meeting Date: 18/Dec/2018

Date Approval Issued: 25/Nov/2018

REB Approval Expiry Date: 22/Dec/2019

Dear Dr. Scott Adams,

The Western University Research Ethics Board has reviewed the application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

Western University REB operates in compliance with, and is constituted in accordance with, the requirements of the TriCouncil Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The REB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Daniel Wyzynski, Research Ethics Coordinator, on behalf of Dr. Joseph Gilbert, HSREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Appendix B: Lawson Health Research Ethics Board approval

LAWSON FINAL APPROVAL NOTICE

LAWSON APPROVAL NUMBER: R-18-193

PROJECT TITLE: The effect of speech rate manipulations on acoustic distinctiveness and speech intelligibility in dysarthria

PRINCIPAL INVESTIGATOR: Dr. Scott Adams

LAWSON APPROVAL DATE: Monday, 16 April 2018

ReDA ID: 4315

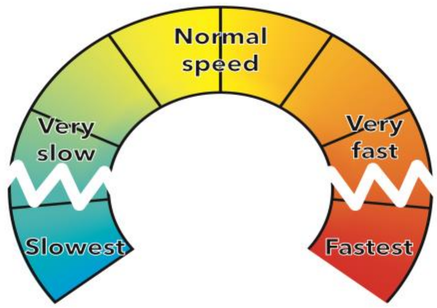
Overall Study Status: Active

Please be advised that the above project was reviewed by Lawson Administration and the project:

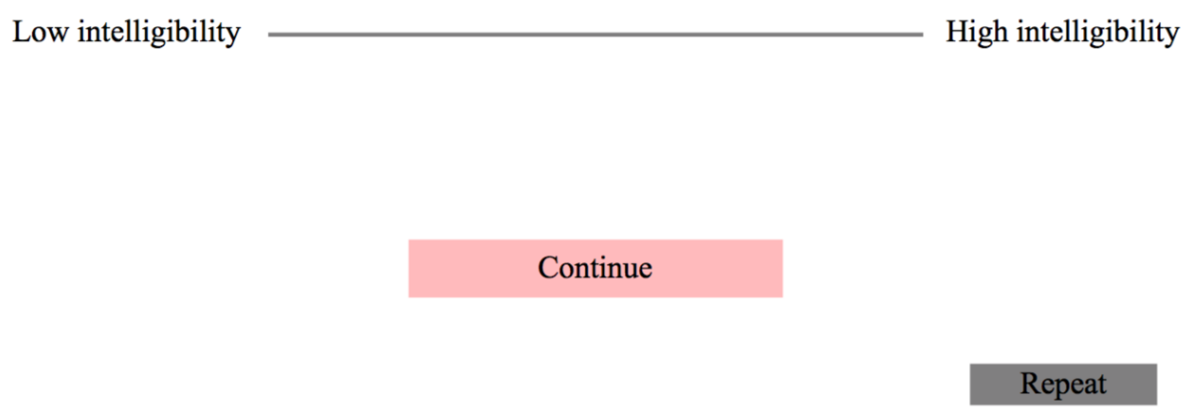
Please provide your Lawson Approval Number (R#) to the appropriate contact(s) in supporting departments (eg. Lab Services, Diagnostic Imaging, etc.) to inform them that your study is starting. The Lawson Approval Number must be provided each time services are requested.

Dr. David Hill
V.P. Research
Lawson Health Research Institute

Appendix C: Visual aid to facilitate speech rate targets



Appendix D: Visual analog scale demonstration



Appendix E: Experiment 1

VOT

Table 10: Summary of fixed effect coefficients for the (log) voice onset time (VOT) model.

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
(INTERCEPT)	-3.367	0.037	69.162	-91.321	0.000
GROUP1	-0.129	0.067	57.952	-1.926	0.059
GROUP2	-0.192	0.078	58.754	-2.449	0.017
RATE.S3	0.309	0.035	86.029	8.869	0.000
RATE.S2	0.223	0.028	80.599	8.081	0.000
RATE.F2	-0.174	0.019	138.142	-8.982	0.000
RATE.F3	-0.337	0.028	118.361	-11.826	0.000
VOICING	-0.684	0.015	537.473	-44.738	0.000
POA1	-0.400	0.019	141.741	-20.858	0.000
POA2	-0.246	0.022	142.854	-11.098	0.000
SEX	0.074	0.034	51.017	2.191	0.033
GROUP1:RATE.S3	0.016	0.066	57.814	0.250	0.803
GROUP2:RATE.S3	0.078	0.079	59.212	0.994	0.324
GROUP1:RATE.S2	-0.048	0.053	53.887	-0.903	0.371
GROUP2:RATE.S2	-0.018	0.061	55.302	-0.294	0.770
GROUP1:RATE.F2	-0.037	0.034	72.129	-1.092	0.279
GROUP2:RATE.F2	0.014	0.041	76.502	0.340	0.735
GROUP1:RATE.F3	-0.088	0.050	61.745	-1.768	0.082
GROUP2:RATE.F3	0.076	0.062	69.965	1.229	0.223
GROUP1:VOICING	-0.045	0.022	7696.086	-2.070	0.038
GROUP2:VOICING	-0.021	0.026	7732.792	-0.798	0.425
VOICING:POA1	-0.079	0.019	141.606	-4.096	0.000
VOICING:POA2	-0.070	0.022	142.725	-3.143	0.002
VOICING:SEX	-0.052	0.005	7655.273	-9.665	0.000
RATE.S3:VOICING	-0.021	0.020	780.735	-1.070	0.285
RATE.S2:VOICING	0.006	0.020	2080.607	0.283	0.777
RATE.F2:VOICING	0.022	0.018	2236.200	1.250	0.211
RATE.F3:VOICING	0.019	0.022	1480.335	0.870	0.385
GROUP1:RATE.S3:VOICING	-0.059	0.029	7656.197	-2.005	0.045
GROUP2:RATE.S3:VOICING	0.012	0.036	7685.507	0.328	0.743

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
GROUP1:RATE.S2:VOICING	-0.023	0.035	7690.714	-0.650	0.516
GROUP2:RATE.S2:VOICING	-0.057	0.041	7718.983	-1.402	0.161
GROUP1:RATE.F2:VOICING	0.079	0.030	7746.834	2.606	0.009
GROUP2:RATE.F2:VOICING	-0.010	0.037	7725.945	-0.269	0.788
GROUP1:RATE.F3:VOICING	0.133	0.034	7709.128	3.911	0.000
GROUP2:RATE.F3:VOICING	0.074	0.043	7651.175	1.707	0.088

Note: GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). POA1 = bilabial vs. alveolar, velar place of articulation; POA2 = alveolar vs. velar place of articulation. Coefficient estimates, standard errors, degrees of freedom (df), t-values, and significances are reported.

Table 11: Summary of post-hoc pairwise comparisons for the three-way interaction between rate, group, and voicing for the (log) voice onset time (VOT) model, demonstrating rate differences.

contrast	group	voicing	estimate	SE	z.ratio	p.value
H1 - S3	OC	voiced	-0.259	0.062	-4.216	0.000
H1 - S2	OC	voiced	-0.182	0.055	-3.276	0.009
H1 - F2	OC	voiced	0.124	0.041	3.058	0.019
H1 - F3	OC	voiced	0.288	0.052	5.527	0.000
S3 - S2	OC	voiced	0.078	0.071	1.101	0.806
S3 - F2	OC	voiced	0.383	0.062	6.151	0.000
S3 - F3	OC	voiced	0.547	0.070	7.861	0.000
S2 - F2	OC	voiced	0.306	0.057	5.362	0.000
S2 - F3	OC	voiced	0.469	0.065	7.228	0.000
F2 - F3	OC	voiced	0.164	0.050	3.300	0.009
H1 - S3	PD	voiced	-0.347	0.055	-6.255	0.000
H1 - S2	PD	voiced	-0.215	0.047	-4.550	0.000
H1 - F2	PD	voiced	0.164	0.037	4.445	0.000
H1 - F3	PD	voiced	0.258	0.048	5.414	0.000
S3 - S2	PD	voiced	0.132	0.061	2.171	0.191
S3 - F2	PD	voiced	0.511	0.056	9.038	0.000
S3 - F3	PD	voiced	0.604	0.063	9.559	0.000
S2 - F2	PD	voiced	0.379	0.049	7.703	0.000
S2 - F3	PD	voiced	0.472	0.057	8.298	0.000
F2 - F3	PD	voiced	0.094	0.045	2.060	0.238
H1 - S3	DBS	voiced	-0.257	0.073	-3.501	0.004
H1 - S2	DBS	voiced	-0.290	0.061	-4.752	0.000
H1 - F2	DBS	voiced	0.168	0.048	3.506	0.004
H1 - F3	DBS	voiced	0.408	0.066	6.148	0.000
S3 - S2	DBS	voiced	-0.033	0.082	-0.405	0.994
S3 - F2	DBS	voiced	0.425	0.075	5.661	0.000
S3 - F3	DBS	voiced	0.664	0.087	7.616	0.000
S2 - F2	DBS	voiced	0.458	0.064	7.207	0.000
S2 - F3	DBS	voiced	0.698	0.078	8.977	0.000
F2 - F3	DBS	voiced	0.240	0.065	3.671	0.002
H1 - S3	OC	voiceless	-0.380	0.061	-6.266	0.000
H1 - S2	OC	voiceless	-0.201	0.054	-3.713	0.002

contrast	group	voicing	estimate	SE	z.ratio	p.value
H1 - F2	OC	voiceless	0.274	0.039	7.026	0.000
H1 - F3	OC	voiceless	0.503	0.051	9.782	0.000
S3 - S2	OC	voiceless	0.180	0.070	2.559	0.078
S3 - F2	OC	voiceless	0.654	0.062	10.574	0.000
S3 - F3	OC	voiceless	0.883	0.070	12.669	0.000
S2 - F2	OC	voiceless	0.474	0.056	8.452	0.000
S2 - F3	OC	voiceless	0.703	0.065	10.863	0.000
F2 - F3	OC	voiceless	0.229	0.049	4.653	0.000
H1 - S3	PD	voiceless	-0.338	0.055	-6.196	0.000
H1 - S2	PD	voiceless	-0.245	0.047	-5.234	0.000
H1 - F2	PD	voiceless	0.146	0.035	4.184	0.000
H1 - F3	PD	voiceless	0.281	0.047	5.971	0.000
S3 - S2	PD	voiceless	0.093	0.060	1.537	0.538
S3 - F2	PD	voiceless	0.484	0.055	8.755	0.000
S3 - F3	PD	voiceless	0.619	0.063	9.840	0.000
S2 - F2	PD	voiceless	0.391	0.048	8.087	0.000
S2 - F3	PD	voiceless	0.526	0.057	9.221	0.000
F2 - F3	PD	voiceless	0.135	0.044	3.046	0.020
H1 - S3	DBS	voiceless	-0.272	0.071	-3.809	0.001
H1 - S2	DBS	voiceless	-0.206	0.060	-3.429	0.005
H1 - F2	DBS	voiceless	0.170	0.043	3.921	0.001
H1 - F3	DBS	voiceless	0.283	0.063	4.462	0.000
S3 - S2	DBS	voiceless	0.066	0.081	0.808	0.928
S3 - F2	DBS	voiceless	0.442	0.073	6.091	0.000
S3 - F3	DBS	voiceless	0.555	0.086	6.485	0.000
S2 - F2	DBS	voiceless	0.376	0.062	6.087	0.000
S2 - F3	DBS	voiceless	0.489	0.077	6.371	0.000
F2 - F3	DBS	voiceless	0.113	0.062	1.827	0.358

Note: These comparisons demonstrate estimated differences between speech rates for each group for voiced and voiceless stops. Results are averaged over PoA, and sex. Results are given on the log scale. Contrasts, group, estimated differences, standard errors, z-ratio, and Tukey-adjusted significances are reported.

Table 12: Summary of post-hoc pairwise comparisons for the three-way interaction between rate, group, and voicing for the (log) voice onset time (VOT) model, demonstrating group differences.

contrast	voicing	prop_wpm_5	estimate	SE	z.ratio	p.value
OC - PD	voiced	H1	-0.068	0.075	-0.909	0.635
OC - DBS	voiced	H1	-0.281	0.089	-3.167	0.004
PD - DBS	voiced	H1	-0.213	0.083	-2.557	0.028
OC - PD	voiceless	H1	0.001	0.075	0.019	1.000
OC - DBS	voiceless	H1	-0.169	0.087	-1.935	0.129
PD - DBS	voiceless	H1	-0.171	0.082	-2.082	0.094
OC - PD	voiced	S3	-0.156	0.096	-1.614	0.240
OC - DBS	voiced	S3	-0.278	0.114	-2.449	0.038
PD - DBS	voiced	S3	-0.123	0.107	-1.145	0.487
OC - PD	voiceless	S3	0.043	0.096	0.450	0.895
OC - DBS	voiceless	S3	-0.061	0.113	-0.539	0.852
PD - DBS	voiceless	S3	-0.104	0.107	-0.978	0.591
OC - PD	voiced	S2	-0.102	0.089	-1.135	0.492
OC - DBS	voiced	S2	-0.389	0.104	-3.737	0.001
PD - DBS	voiced	S2	-0.288	0.097	-2.976	0.008
OC - PD	voiceless	S2	-0.044	0.089	-0.487	0.877
OC - DBS	voiceless	S2	-0.175	0.104	-1.682	0.212
PD - DBS	voiceless	S2	-0.131	0.097	-1.357	0.363
OC - PD	voiced	F2	-0.028	0.076	-0.373	0.926
OC - DBS	voiced	F2	-0.237	0.090	-2.633	0.023
PD - DBS	voiced	F2	-0.208	0.085	-2.460	0.037
OC - PD	voiceless	F2	-0.126	0.075	-1.674	0.215
OC - DBS	voiceless	F2	-0.273	0.088	-3.091	0.006
PD - DBS	voiceless	F2	-0.147	0.083	-1.774	0.178
OC - PD	voiced	F3	-0.099	0.086	-1.148	0.485
OC - DBS	voiced	F3	-0.161	0.104	-1.549	0.268
PD - DBS	voiced	F3	-0.062	0.099	-0.632	0.803
OC - PD	voiceless	F3	-0.221	0.086	-2.563	0.028
OC - DBS	voiceless	F3	-0.389	0.103	-3.761	0.000
PD - DBS	voiceless	F3	-0.168	0.098	-1.717	0.199

contrast	voicing	prop_wpm_5	estimate	SE	z.ratio	p.value

Note: These comparisons demonstrate estimated differences between groups, for each speech rate, for voiced and voiceless stops. Results are averaged over PoA, and sex. Results are given on the log scale. Contrasts, group, estimated differences, standard errors, z-ratio, and Tukey-adjusted significances are reported.

VDC

Table 13: Summary of fixed effect coefficients for the dichotomized voicing during closure (VDC) model.

Contrast	Estimate	Std. Error	z value	Pr(>|z|)
(INTERCEPT)	-3.159	0.436	-7.239	0.000
GROUP1	-1.104	0.782	-1.411	0.158
GROUP2	-0.114	0.904	-0.126	0.900
RATE.S3	-1.046	0.368	-2.838	0.005
RATE.S2	-0.553	0.353	-1.567	0.117
RATE.F2	1.464	0.267	5.478	0.000
RATE.F3	2.730	0.305	8.962	0.000
VOICING	3.112	0.207	15.043	0.000
POA1	-0.562	0.206	-2.725	0.006
POA2	-0.874	0.239	-3.656	0.000
GROUP1:RATE.S3	0.365	0.517	0.706	0.480
GROUP2:RATE.S3	0.417	0.617	0.676	0.499
GROUP1:RATE.S2	0.242	0.459	0.526	0.599
GROUP2:RATE.S2	0.544	0.497	1.096	0.273
GROUP1:RATE.F2	0.230	0.420	0.548	0.583
GROUP2:RATE.F2	0.145	0.507	0.286	0.775
GROUP1:RATE.F3	0.674	0.519	1.299	0.194
GROUP2:RATE.F3	0.189	0.630	0.300	0.764
GROUP1:VOICING	0.488	0.179	2.724	0.006
GROUP2:VOICING	0.620	0.179	3.462	0.001
RATE.S3:VOICING	-0.443	0.303	-1.461	0.144
RATE.S2:VOICING	-0.427	0.324	-1.319	0.187
RATE.F2:VOICING	-0.420	0.240	-1.748	0.080
RATE.F3:VOICING	-0.908	0.237	-3.838	0.000
RATE.S3:POA1	0.572	0.271	2.110	0.035
RATE.S2:POA1	0.613	0.318	1.930	0.054
RATE.F2:POA1	0.212	0.275	0.770	0.441
RATE.F3:POA1	-0.255	0.290	-0.877	0.381
RATE.S3:POA2	0.386	0.314	1.230	0.219
RATE.S2:POA2	0.155	0.367	0.421	0.673
RATE.F2:POA2	0.368	0.319	1.156	0.248

Contrast	Estimate	Std. Error	z value	Pr(>|z|)
RATE.F3:POA2	1.093	0.346	3.161	0.002
VOICED:SEX	0.191	0.388	0.492	0.623
VOICELESS:SEX	0.719	0.414	1.737	0.082

Note: GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). POA1 = bilabial vs. alveolar, velar place of articulation; POA2 = alveolar vs. velar place of articulation. VDC is dichotomized as Total VDC vs Some or No VDC. Coefficient estimates, standard errors, z-values, and significances are reported.

QVAI

Table 14: Summary of fixed effect coefficients for the quadrilateral vowel articulation index (QVAI) model.

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
(INTERCEPT)	1.633	0.019	57.213	84.235	0.000
GROUP1	0.076	0.037	57.213	2.069	0.043
GROUP2	0.088	0.043	57.213	2.051	0.045
RATE.S3	0.038	0.019	50.777	1.998	0.051
RATE.S2	0.022	0.014	38.399	1.600	0.118
RATE.F2	-0.077	0.011	57.315	-6.891	0.000
RATE.F3	-0.128	0.015	55.643	-8.301	0.000
SEX	0.053	0.019	57.213	2.783	0.007
GROUP1:RATE.S3	-0.021	0.036	50.777	-0.579	0.565
GROUP2:RATE.S3	0.012	0.042	50.777	0.298	0.767
GROUP1:RATE.S2	-0.041	0.026	38.215	-1.539	0.132
GROUP2:RATE.S2	0.044	0.030	37.802	1.501	0.142
GROUP1:RATE.F2	0.021	0.021	57.315	0.991	0.326
GROUP2:RATE.F2	-0.011	0.025	57.315	-0.458	0.649
GROUP1:RATE.F3	0.025	0.029	55.936	0.856	0.395
GROUP2:RATE.F3	-0.059	0.035	55.345	-1.704	0.094
RATE.S3:SEX	0.032	0.019	50.777	1.721	0.091
RATE.S2:SEX	0.018	0.014	38.479	1.303	0.200
RATE.F2:SEX	-0.027	0.011	57.315	-2.404	0.019
RATE.F3:SEX	-0.040	0.015	56.080	-2.638	0.011

Note: GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). Coefficient estimates, standard errors, degrees of freedom (df), t-values, and significances.

Intensity

Table 15: Summary of fixed effect coefficients for the vowel intensity model.

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
(INTERCEPT)	73.936	0.554	55.712	133.386	0.000
GROUP1	1.471	1.055	52.371	1.394	0.169
GROUP2	1.275	1.448	52.732	0.881	0.383
RATE.S3	-2.798	0.153	678.547	-18.263	0.000
RATE.S2	-2.042	0.208	91.368	-9.816	0.000
RATE.F2	0.168	0.417	58.236	0.404	0.688
RATE.F3	0.511	0.471	57.880	1.086	0.282
SEX	-1.365	0.546	52.592	-2.498	0.016
GROUP1:RATE.S3	1.486	0.193	7648.349	7.682	0.000
GROUP2:RATE.S3	1.709	0.278	7633.083	6.153	0.000
GROUP1:RATE.S2	0.687	0.373	62.078	1.844	0.070
GROUP2:RATE.S2	0.939	0.493	63.343	1.903	0.062
GROUP1:RATE.F2	-0.006	0.790	53.820	-0.008	0.994
GROUP2:RATE.F2	-0.005	1.087	54.831	-0.005	0.996
GROUP1:RATE.F3	0.840	0.887	52.417	0.946	0.348
GROUP2:RATE.F3	0.428	1.217	52.731	0.352	0.727
GROUP1:SEX	0.489	1.055	52.382	0.464	0.645
GROUP2:SEX	0.710	1.448	52.740	0.490	0.626
RATE.S3:SEX	-0.270	0.103	7662.241	-2.620	0.009
RATE.S2:SEX	-0.105	0.189	63.093	-0.556	0.580
RATE.F2:SEX	-0.301	0.410	54.424	-0.734	0.466
RATE.F3:SEX	-0.490	0.459	52.570	-1.067	0.291
GROUP1:RATE.S3:SEX	2.163	0.194	7664.550	11.169	0.000
GROUP2:RATE.S3:SEX	0.835	0.278	7661.034	3.002	0.003
GROUP1:RATE.S2:SEX	1.032	0.373	62.268	2.766	0.007
GROUP2:RATE.S2:SEX	0.971	0.493	63.557	1.968	0.053
GROUP1:RATE.F2:SEX	-0.293	0.790	53.824	-0.371	0.712
GROUP2:RATE.F2:SEX	1.224	1.087	54.856	1.126	0.265
GROUP1:RATE.F3:SEX	-0.391	0.887	52.371	-0.441	0.661
GROUP2:RATE.F3:SEX	2.124	1.217	52.711	1.744	0.087

Note: GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). Coefficient estimates, standard errors, degrees of freedom (df), t-values, and significances are reported.

Table 16: Summary of post-hoc pairwise comparisons for the two-way interaction between group and rate for the speech intensity model.

contrast	group	estimate	SE	z.ratio	p.value
H1 - S3	OC	1.808	0.177	10.229	0.000
H1 - S2	OC	1.584	0.293	5.406	0.000
H1 - F2	OC	-0.164	0.579	-0.284	0.999
H1 - F3	OC	-1.071	0.655	-1.634	0.475
S3 - S2	OC	-0.224	0.285	-0.785	0.935
S3 - F2	OC	-1.972	0.585	-3.371	0.007
S3 - F3	OC	-2.878	0.659	-4.367	0.000
S2 - F2	OC	-1.748	0.632	-2.768	0.045
S2 - F3	OC	-2.654	0.701	-3.786	0.001
F2 - F3	OC	-0.906	0.854	-1.061	0.826
H1 - S3	PD	2.440	0.181	13.463	0.000
H1 - S2	PD	1.802	0.289	6.234	0.000
H1 - F2	PD	-0.168	0.597	-0.281	0.999
H1 - F3	PD	-0.445	0.674	-0.660	0.965
S3 - S2	PD	-0.638	0.277	-2.301	0.145
S3 - F2	PD	-2.608	0.603	-4.326	0.000
S3 - F3	PD	-2.885	0.678	-4.254	0.000
S2 - F2	PD	-1.969	0.644	-3.057	0.019
S2 - F3	PD	-2.246	0.715	-3.140	0.015
F2 - F3	PD	-0.277	0.880	-0.315	0.998
H1 - S3	DBS	4.148	0.264	15.692	0.000
H1 - S2	DBS	2.740	0.418	6.557	0.000
H1 - F2	DBS	-0.173	0.914	-0.189	1.000
H1 - F3	DBS	-0.017	1.023	-0.017	1.000
S3 - S2	DBS	-1.408	0.394	-3.571	0.003
S3 - F2	DBS	-4.321	0.910	-4.749	0.000
S3 - F3	DBS	-4.165	1.019	-4.088	0.000
S2 - F2	DBS	-2.913	0.966	-3.016	0.022
S2 - F3	DBS	-2.757	1.069	-2.578	0.074
F2 - F3	DBS	0.156	1.340	0.117	1.000

contrast	group	estimate	SE	z.ratio	p.value

Note: These comparisons demonstrate estimated differences between groups for each speech rate. Results are averaged over sex. Contrasts, group, estimated differences, standard errors, z-ratio, and Tukey-adjusted significances are reported.

HNR

Table 17: Summary of all fixed effects coefficients for the harmonics-to-noise (HNR) model.

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
(INTERCEPT)	20.430	0.532	56.198	38.434	0.000
GROUP1	0.231	1.013	53.050	0.228	0.821
GROUP2	-0.569	1.393	53.879	-0.408	0.685
RATE.S3	-1.106	0.429	61.552	-2.581	0.012
RATE.S2	-0.602	0.341	57.714	-1.766	0.083
RATE.F2	-0.657	0.289	69.119	-2.274	0.026
RATE.F3	-1.324	0.374	58.952	-3.538	0.001
V.HEIGHT	3.152	0.117	620.719	27.033	0.000
SEX	0.039	0.525	53.550	0.075	0.941
GROUP1:RATE.S3	2.330	0.798	53.117	2.920	0.005
GROUP2:RATE.S3	4.042	1.101	54.523	3.673	0.001
GROUP1:RATE.S2	1.412	0.650	49.907	2.174	0.034
GROUP2:RATE.S2	1.495	0.857	49.937	1.745	0.087
GROUP1:RATE.F2	-0.626	0.528	55.668	-1.184	0.241
GROUP2:RATE.F2	0.161	0.740	60.899	0.217	0.829
GROUP1:RATE.F3	-0.214	0.691	49.102	-0.310	0.758
GROUP2:RATE.F3	0.032	0.952	50.398	0.034	0.973
GROUP1:V.HEIGHT	0.284	0.180	7958.492	1.577	0.115
GROUP2:V.HEIGHT	-0.142	0.214	7997.654	-0.663	0.507
RATE.S3:V.HEIGHT	-0.179	0.153	743.214	-1.167	0.244
RATE.S2:V.HEIGHT	-0.262	0.159	2033.305	-1.647	0.100
RATE.F2:V.HEIGHT	-0.255	0.140	2055.622	-1.828	0.068
RATE.F3:V.HEIGHT	-0.555	0.167	1318.720	-3.323	0.001
RATE.S3:SEX	-1.677	0.414	53.966	-4.045	0.000
RATE.S2:SEX	-0.925	0.329	50.046	-2.811	0.007
RATE.F2:SEX	-0.554	0.277	58.739	-2.002	0.050
RATE.F3:SEX	-0.891	0.359	49.785	-2.484	0.016
V.HEIGHT:SEX	0.185	0.044	7914.765	4.201	0.000
GROUP1:SEX	-0.535	1.013	53.063	-0.528	0.600
GROUP2:SEX	-2.619	1.393	53.890	-1.880	0.066
GROUP1:RATE.S3:V.HEIGHT	-0.615	0.240	7928.117	-2.563	0.010

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
GROUP2:RATE.S3:V.HEIGHT	-0.293	0.292	7959.768	-1.006	0.315
GROUP1:RATE.S2:V.HEIGHT	-0.718	0.288	7972.754	-2.496	0.013
GROUP2:RATE.S2:V.HEIGHT	0.079	0.333	7989.870	0.236	0.814
GROUP1:RATE.F2:V.HEIGHT	-0.564	0.248	7997.710	-2.276	0.023
GROUP2:RATE.F2:V.HEIGHT	-0.833	0.297	7991.016	-2.803	0.005
GROUP1:RATE.F3:V.HEIGHT	-0.464	0.275	8009.343	-1.688	0.092
GROUP2:RATE.F3:V.HEIGHT	-0.373	0.343	8040.567	-1.088	0.276
GROUP1:RATE.S3:SEX	1.758	0.798	53.135	2.203	0.032
GROUP2:RATE.S3:SEX	2.502	1.101	54.554	2.273	0.027
GROUP1:RATE.S2:SEX	0.400	0.650	49.974	0.616	0.541
GROUP2:RATE.S2:SEX	0.496	0.857	49.982	0.579	0.565
GROUP1:RATE.F2:SEX	0.320	0.528	55.686	0.605	0.548
GROUP2:RATE.F2:SEX	1.874	0.740	60.967	2.532	0.014
GROUP1:RATE.F3:SEX	0.719	0.691	48.987	1.041	0.303
GROUP2:RATE.F3:SEX	2.496	0.952	50.324	2.621	0.012

Note: GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). POA1 = bilabial vs. alveolar, velar place of articulation; POA2 = alveolar vs. velar place of articulation. Coefficient estimates, standard errors, degrees of freedom (df), t-values, and significances are reported.

Table 18: Summary of post-hoc pairwise comparisons for the three-way interaction between group, rate and vowel height for the harmonics-to-noise (HNR) model.

contrast	group	vwl_height	estimate	SE	z.ratio	p.value
H1 - S3	OC	high	0.141	0.627	0.226	0.999
H1 - S2	OC	high	0.401	0.560	0.715	0.953
H1 - F2	OC	high	1.705	0.443	3.852	0.001
H1 - F3	OC	high	2.331	0.569	4.100	0.000
S3 - S2	OC	high	0.259	0.770	0.337	0.997
S3 - F2	OC	high	1.564	0.699	2.237	0.166
S3 - F3	OC	high	2.190	0.781	2.802	0.041
S2 - F2	OC	high	1.304	0.644	2.025	0.254
S2 - F3	OC	high	1.930	0.733	2.633	0.065
F2 - F3	OC	high	0.626	0.634	0.988	0.861
H1 - S3	PD	high	-0.018	0.640	-0.028	1.000
H1 - S2	PD	high	0.308	0.530	0.582	0.978
H1 - F2	PD	high	0.852	0.445	1.914	0.310
H1 - F3	PD	high	1.823	0.575	3.171	0.013
S3 - S2	PD	high	0.326	0.760	0.429	0.993
S3 - F2	PD	high	0.870	0.715	1.216	0.742
S3 - F3	PD	high	1.841	0.799	2.303	0.144
S2 - F2	PD	high	0.543	0.621	0.875	0.906
S2 - F3	PD	high	1.514	0.717	2.113	0.214
F2 - F3	PD	high	0.971	0.645	1.506	0.559
H1 - S3	DBS	high	3.732	0.963	3.876	0.001
H1 - S2	DBS	high	1.882	0.772	2.437	0.106
H1 - F2	DBS	high	0.180	0.672	0.267	0.999
H1 - F3	DBS	high	1.482	0.857	1.729	0.416
S3 - S2	DBS	high	-1.849	1.133	-1.632	0.476
S3 - F2	DBS	high	-3.552	1.074	-3.307	0.008
S3 - F3	DBS	high	-2.250	1.197	-1.880	0.328
S2 - F2	DBS	high	-1.702	0.909	-1.873	0.332
S2 - F3	DBS	high	-0.400	1.052	-0.380	0.996
F2 - F3	DBS	high	1.302	0.973	1.338	0.667
H1 - S3	OC	low	-1.036	0.628	-1.651	0.465
H1 - S2	OC	low	-1.079	0.553	-1.954	0.289
H1 - F2	OC	low	0.443	0.446	0.994	0.858

contrast	group	vwl_height	estimate	SE	z.ratio	p.value
H1 - F3	OC	low	0.603	0.568	1.062	0.826
S3 - S2	OC	low	-0.043	0.765	-0.056	1.000
S3 - F2	OC	low	1.479	0.701	2.112	0.215
S3 - F3	OC	low	1.639	0.780	2.100	0.220
S2 - F2	OC	low	1.522	0.639	2.384	0.120
S2 - F3	OC	low	1.682	0.726	2.318	0.139
F2 - F3	OC	low	0.160	0.635	0.252	0.999
H1 - S3	PD	low	-0.259	0.640	-0.405	0.994
H1 - S2	PD	low	0.343	0.534	0.641	0.968
H1 - F2	PD	low	-0.116	0.447	-0.259	0.999
H1 - F3	PD	low	0.650	0.575	1.131	0.790
S3 - S2	PD	low	0.602	0.761	0.791	0.933
S3 - F2	PD	low	0.143	0.715	0.200	1.000
S3 - F3	PD	low	0.909	0.798	1.140	0.785
S2 - F2	PD	low	-0.458	0.625	-0.734	0.949
S2 - F3	PD	low	0.307	0.718	0.428	0.993
F2 - F3	PD	low	0.766	0.645	1.187	0.759
H1 - S3	DBS	low	4.077	0.966	4.221	0.000
H1 - S2	DBS	low	1.759	0.765	2.299	0.145
H1 - F2	DBS	low	0.878	0.681	1.289	0.698
H1 - F3	DBS	low	1.056	0.854	1.236	0.730
S3 - S2	DBS	low	-2.317	1.126	-2.058	0.238
S3 - F2	DBS	low	-3.199	1.077	-2.969	0.025
S3 - F3	DBS	low	-3.021	1.193	-2.533	0.083
S2 - F2	DBS	low	-0.882	0.904	-0.976	0.866
S2 - F3	DBS	low	-0.703	1.039	-0.677	0.961
F2 - F3	DBS	low	0.178	0.971	0.183	1.000

Note: These comparisons demonstrate estimated differences between speech rates for each group for high and low vowels. Results are averaged over sex. Contrasts, group, estimated differences, standard errors, z-ratio, and Tukey-adjusted significances are reported.

Appendix F: Experiment 2

Stop consonant accuracy

Table 19: Summary of fixed effect coefficients for the stop consonant accuracy model.

Contrast	Estimate	Std. Error	z value	Pr(>|z|)
(INTERCEPT)	1.126	0.119	9.445	0.000
GROUP0	1.671	0.219	7.648	0.000
GROUP1	0.984	0.221	4.444	0.000
GROUP2	0.089	0.262	0.339	0.735
RATE.S3	-0.282	0.109	-2.598	0.009
RATE.S2	-0.222	0.095	-2.348	0.019
RATE.F2	-0.046	0.079	-0.576	0.564
RATE.F3	-0.065	0.127	-0.513	0.608
LISTENER2	-0.338	0.078	-4.325	0.000
LISTENER3	-0.662	0.081	-8.138	0.000
LISTENER4	-0.358	0.076	-4.718	0.000
LISTENER5	-0.896	0.075	-12.019	0.000
LISTENER6	-0.757	0.079	-9.624	0.000
LISTENER7	-0.392	0.085	-4.613	0.000
LISTENER8	-1.470	0.140	-10.500	0.000
VOICING	-0.202	0.066	-3.041	0.002
POA1	-0.007	0.097	-0.071	0.944
POA2	0.468	0.112	4.165	0.000
GROUP0:RATE.S3	-0.176	0.199	-0.883	0.377
GROUP1:RATE.S3	-0.354	0.199	-1.774	0.076
GROUP2:RATE.S3	-0.055	0.237	-0.231	0.817
GROUP0:RATE.S2	-0.284	0.193	-1.478	0.139
GROUP1:RATE.S2	-0.141	0.180	-0.787	0.431
GROUP2:RATE.S2	0.130	0.202	0.642	0.521
GROUP0:RATE.F2	-0.218	0.164	-1.331	0.183
GROUP1:RATE.F2	0.119	0.150	0.798	0.425
GROUP2:RATE.F2	0.004	0.172	0.025	0.980
GROUP0:RATE.F3	-0.564	0.258	-2.181	0.029
GROUP1:RATE.F3	0.394	0.267	1.476	0.140
GROUP2:RATE.F3	0.124	0.319	0.389	0.697
GROUP0:VOICING	0.211	0.108	1.960	0.050

Contrast	Estimate	Std. Error	z value	Pr(>|z|)
GROUP1:VOICING	-0.160	0.089	-1.800	0.072
GROUP2:VOICING	-0.158	0.102	-1.550	0.121
GROUP0:POA1	-0.502	0.088	-5.694	0.000
GROUP1:POA1	0.189	0.087	2.165	0.030
GROUP2:POA1	0.297	0.104	2.865	0.004
GROUP0:POA2	0.445	0.107	4.151	0.000
GROUP1:POA2	0.293	0.101	2.907	0.004
GROUP2:POA2	0.247	0.121	2.047	0.041
RATE.S3:VOICING	-0.023	0.081	-0.289	0.773
RATE.S2:VOICING	-0.041	0.079	-0.524	0.601
RATE.F2:VOICING	-0.054	0.069	-0.780	0.435
RATE.F3:VOICING	0.006	0.083	0.076	0.939
GROUP0:RATE.S3:VOICING	-0.222	0.129	-1.719	0.086
GROUP1:RATE.S3:VOICING	-0.044	0.117	-0.374	0.708
GROUP2:RATE.S3:VOICING	0.239	0.139	1.721	0.085
GROUP0:RATE.S2:VOICING	0.246	0.158	1.552	0.121
GROUP1:RATE.S2:VOICING	-0.007	0.137	-0.053	0.957
GROUP2:RATE.S2:VOICING	0.290	0.156	1.855	0.064
GROUP0:RATE.F2:VOICING	-0.166	0.141	-1.184	0.236
GROUP1:RATE.F2:VOICING	0.191	0.121	1.575	0.115
GROUP2:RATE.F2:VOICING	0.036	0.137	0.264	0.792
GROUP0:RATE.F3:VOICING	-0.167	0.146	-1.140	0.254
GROUP1:RATE.F3:VOICING	0.082	0.139	0.594	0.553
GROUP2:RATE.F3:VOICING	0.095	0.163	0.582	0.561

Note: GROUP0 YC vs. OC, PD, DBS; GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). POA1 = bilabial vs. alveolar, velar place of articulation; POA2 = alveolar vs. velar place of articulation. Coefficient estimates, standard errors, z-values, and significances are reported.

Table 20: Summary of post-hoc pairwise comparisons for the two-way interaction between group and speech rate for the stop consonant accuracy model.

contrast	group	estimate	SE	z.ratio	p.value
H1 - S3	YC	0.414	0.190	2.176	0.189
H1 - S2	YC	0.435	0.185	2.347	0.130
H1 - F2	YC	0.209	0.158	1.328	0.674
H1 - F3	YC	0.488	0.235	2.074	0.231
S3 - S2	YC	0.021	0.206	0.103	1.000
S3 - F2	YC	-0.205	0.196	-1.042	0.836
S3 - F3	YC	0.074	0.261	0.284	0.999
S2 - F2	YC	-0.226	0.194	-1.168	0.770
S2 - F3	YC	0.053	0.259	0.203	1.000
F2 - F3	YC	0.279	0.233	1.199	0.752
H1 - S3	OC	0.474	0.175	2.706	0.053
H1 - S2	OC	0.245	0.159	1.538	0.537
H1 - F2	OC	-0.089	0.133	-0.667	0.963
H1 - F3	OC	-0.338	0.225	-1.505	0.559
S3 - S2	OC	-0.229	0.196	-1.169	0.769
S3 - F2	OC	-0.562	0.189	-2.974	0.025
S3 - F3	OC	-0.812	0.259	-3.129	0.015
S2 - F2	OC	-0.334	0.177	-1.882	0.327
S2 - F3	OC	-0.583	0.251	-2.322	0.138
F2 - F3	OC	-0.250	0.226	-1.106	0.803
H1 - S3	PD	0.148	0.157	0.937	0.882
H1 - S2	PD	0.039	0.135	0.286	0.999
H1 - F2	PD	0.029	0.113	0.255	0.999
H1 - F3	PD	-0.006	0.196	-0.033	1.000
S3 - S2	PD	-0.109	0.168	-0.650	0.967
S3 - F2	PD	-0.119	0.170	-0.701	0.956
S3 - F3	PD	-0.154	0.231	-0.667	0.963
S2 - F2	PD	-0.010	0.151	-0.066	1.000
S2 - F3	PD	-0.045	0.218	-0.206	1.000
F2 - F3	PD	-0.035	0.195	-0.180	1.000
H1 - S3	DBS	0.093	0.199	0.465	0.990
H1 - S2	DBS	0.169	0.166	1.016	0.848
H1 - F2	DBS	0.033	0.140	0.235	0.999

contrast	group	estimate	SE	z.ratio	p.value
H1 - F3	DBS	0.118	0.263	0.447	0.992
S3 - S2	DBS	0.076	0.222	0.342	0.997
S3 - F2	DBS	-0.060	0.217	-0.277	0.999
S3 - F3	DBS	0.025	0.309	0.081	1.000
S2 - F2	DBS	-0.136	0.187	-0.726	0.950
S2 - F3	DBS	-0.051	0.289	-0.176	1.000
F2 - F3	DBS	0.085	0.268	0.317	0.998

Note: These comparisons demonstrate estimated differences between speech rates for each group. Results are averaged over listener, voicing, and PoA. Results are given on the log odds ratio scale. Contrast, group, estimated difference, standard errors, z-ratio, and Tukey-adjusted significances are reported.

Vowel consonant accuracy

Table 21: Summary of fixed effect coefficients for the vowel accuracy model.

Contrast	Estimate	Std. Error	z value	Pr(>|z|)
(INTERCEPT)	1.867	0.111	16.879	0.000
GROUP0	0.905	0.214	4.228	0.000
GROUP1	1.011	0.220	4.590	0.000
GROUP2	0.337	0.248	1.357	0.175
RATE.S3	-0.268	0.094	-2.849	0.004
RATE.S2	-0.253	0.095	-2.658	0.008
RATE.F2	-0.047	0.094	-0.495	0.620
RATE.F3	-0.216	0.133	-1.621	0.105
LISTENER2	-0.715	0.083	-8.612	0.000
LISTENER3	-0.551	0.088	-6.278	0.000
LISTENER4	-0.364	0.080	-4.536	0.000
LISTENER5	-0.592	0.084	-7.032	0.000
LISTENER6	-1.182	0.083	-14.157	0.000
LISTENER7	-0.663	0.090	-7.329	0.000
LISTENER8	-0.135	0.124	-1.089	0.276
V.BACKNESS	-0.607	0.055	-11.046	0.000
GROUP0:RATE.S3	0.507	0.206	2.457	0.014
GROUP1:RATE.S3	-0.536	0.207	-2.587	0.010
GROUP2:RATE.S3	-0.357	0.225	-1.587	0.113
GROUP0:RATE.S2	0.471	0.229	2.058	0.040
GROUP1:RATE.S2	-0.354	0.212	-1.665	0.096
GROUP2:RATE.S2	-0.291	0.218	-1.330	0.183
GROUP0:RATE.F2	0.381	0.227	1.681	0.093
GROUP1:RATE.F2	-0.241	0.212	-1.135	0.256
GROUP2:RATE.F2	-0.583	0.230	-2.532	0.011
GROUP0:RATE.F3	-0.158	0.292	-0.541	0.589
GROUP1:RATE.F3	0.106	0.311	0.341	0.733
GROUP2:RATE.F3	-0.672	0.355	-1.897	0.058
GROUP0:V.BACKNESS	-0.101	0.122	-0.827	0.408
GROUP1:V.BACKNESS	0.072	0.114	0.634	0.526
GROUP2:V.BACKNESS	-0.313	0.108	-2.892	0.004
RATE.S3:V.BACKNESS	0.038	0.071	0.527	0.598
RATE.S2:V.BACKNESS	0.218	0.076	2.870	0.004

Contrast	Estimate	Std. Error	z value	Pr(>|z|)
RATE.F2:V.BACKNESS	-0.478	0.071	-6.693	0.000
RATE.F3:V.BACKNESS	-0.510	0.081	-6.291	0.000
GROUP0:RATE.S3:V.BACKNESS	-0.180	0.154	-1.170	0.242
GROUP1:RATE.S3:V.BACKNESS	-0.329	0.146	-2.249	0.025
GROUP2:RATE.S3:V.BACKNESS	0.395	0.145	2.726	0.006
GROUP0:RATE.S2:V.BACKNESS	-0.120	0.189	-0.637	0.524
GROUP1:RATE.S2:V.BACKNESS	-0.400	0.160	-2.496	0.013
GROUP2:RATE.S2:V.BACKNESS	0.308	0.158	1.945	0.052
GROUP0:RATE.F2:V.BACKNESS	-0.259	0.177	-1.462	0.144
GROUP1:RATE.F2:V.BACKNESS	-0.326	0.154	-2.113	0.035
GROUP2:RATE.F2:V.BACKNESS	0.049	0.151	0.323	0.747
GROUP0:RATE.F3:V.BACKNESS	0.016	0.177	0.091	0.928
GROUP1:RATE.F3:V.BACKNESS	-0.230	0.178	-1.291	0.197
GROUP2:RATE.F3:V.BACKNESS	0.299	0.179	1.673	0.094

Note: GROUP0 YC vs. OC, PD, DBS; GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). V.BACK = Vowel backness. Coefficient estimates, standard errors, z-values, and significances are reported.

Table 22: Summary of post-hoc pairwise comparisons for the three-way interaction between group, speech rate, and vowel backness for the vowel accuracy model.

contrast	group	vwl_back	estimate	SE	z.ratio	p.value
H1 - S3	YC	front	-0.015	0.208	-0.070	1.000
H1 - S2	YC	front	-0.228	0.240	-0.953	0.876
H1 - F2	YC	front	0.432	0.209	2.069	0.233
H1 - F3	YC	front	0.832	0.277	3.007	0.022
S3 - S2	YC	front	-0.214	0.247	-0.865	0.910
S3 - F2	YC	front	0.447	0.228	1.961	0.285
S3 - F3	YC	front	0.847	0.288	2.936	0.028
S2 - F2	YC	front	0.661	0.259	2.548	0.080
S2 - F3	YC	front	1.061	0.314	3.373	0.007
F2 - F3	YC	front	0.400	0.285	1.402	0.626
H1 - S3	OC	front	0.889	0.200	4.438	0.000
H1 - S2	OC	front	0.625	0.212	2.949	0.026
H1 - F2	OC	front	0.933	0.196	4.755	0.000
H1 - F3	OC	front	0.773	0.274	2.824	0.038
S3 - S2	OC	front	-0.264	0.225	-1.174	0.767
S3 - F2	OC	front	0.044	0.223	0.199	1.000
S3 - F3	OC	front	-0.116	0.289	-0.401	0.995
S2 - F2	OC	front	0.309	0.238	1.297	0.693
S2 - F3	OC	front	0.148	0.301	0.492	0.988
F2 - F3	OC	front	-0.160	0.279	-0.574	0.979
H1 - S3	PD	front	0.005	0.167	0.030	1.000
H1 - S2	PD	front	-0.138	0.169	-0.817	0.926
H1 - F2	PD	front	0.633	0.165	3.827	0.001
H1 - F3	PD	front	0.835	0.237	3.521	0.004
S3 - S2	PD	front	-0.143	0.188	-0.762	0.941
S3 - F2	PD	front	0.628	0.197	3.189	0.012
S3 - F3	PD	front	0.830	0.257	3.229	0.011
S2 - F2	PD	front	0.771	0.203	3.806	0.001
S2 - F3	PD	front	0.973	0.262	3.716	0.002
F2 - F3	PD	front	0.202	0.249	0.813	0.927
H1 - S3	DBS	front	0.043	0.216	0.200	1.000
H1 - S2	DBS	front	-0.120	0.212	-0.569	0.980
H1 - F2	DBS	front	0.098	0.208	0.472	0.990

contrast	group	vwl_back	estimate	SE	z.ratio	p.value
H1 - F3	DBS	front	0.462	0.319	1.447	0.597
S3 - S2	DBS	front	-0.164	0.248	-0.659	0.965
S3 - F2	DBS	front	0.055	0.253	0.217	1.000
S3 - F3	DBS	front	0.419	0.348	1.204	0.749
S2 - F2	DBS	front	0.219	0.252	0.867	0.909
S2 - F3	DBS	front	0.582	0.348	1.675	0.450
F2 - F3	DBS	front	0.364	0.338	1.076	0.819
H1 - S3	YC	back	-0.209	0.268	-0.782	0.936
H1 - S2	YC	back	0.028	0.314	0.089	1.000
H1 - F2	YC	back	-0.911	0.317	-2.876	0.033
H1 - F3	YC	back	-0.163	0.336	-0.487	0.989
S3 - S2	YC	back	0.237	0.307	0.773	0.939
S3 - F2	YC	back	-0.701	0.318	-2.206	0.177
S3 - F3	YC	back	0.046	0.333	0.138	1.000
S2 - F2	YC	back	-0.939	0.359	-2.613	0.068
S2 - F3	YC	back	-0.191	0.374	-0.511	0.986
F2 - F3	YC	back	0.747	0.373	2.006	0.263
H1 - S3	OC	back	0.616	0.232	2.651	0.062
H1 - S2	OC	back	0.588	0.248	2.374	0.122
H1 - F2	OC	back	-0.328	0.252	-1.301	0.690
H1 - F3	OC	back	-0.562	0.326	-1.722	0.420
S3 - S2	OC	back	-0.028	0.248	-0.112	1.000
S3 - F2	OC	back	-0.943	0.262	-3.603	0.003
S3 - F3	OC	back	-1.177	0.330	-3.566	0.003
S2 - F2	OC	back	-0.916	0.280	-3.265	0.010
S2 - F3	OC	back	-1.150	0.346	-3.326	0.008
F2 - F3	OC	back	-0.234	0.338	-0.692	0.958
H1 - S3	PD	back	0.785	0.185	4.232	0.000
H1 - S2	PD	back	0.934	0.184	5.084	0.000
H1 - F2	PD	back	0.073	0.186	0.395	0.995
H1 - F3	PD	back	0.260	0.250	1.042	0.836
S3 - S2	PD	back	0.149	0.190	0.784	0.935
S3 - F2	PD	back	-0.711	0.206	-3.450	0.005
S3 - F3	PD	back	-0.525	0.261	-2.007	0.262
S2 - F2	PD	back	-0.860	0.208	-4.132	0.000

contrast	group	vwl_back	estimate	SE	z.ratio	p.value
S2 - F3	PD	back	-0.673	0.264	-2.555	0.079
F2 - F3	PD	back	0.187	0.256	0.729	0.950
H1 - S3	DBS	back	0.033	0.224	0.146	1.000
H1 - S2	DBS	back	0.335	0.220	1.523	0.547
H1 - F2	DBS	back	-0.559	0.229	-2.441	0.105
H1 - F3	DBS	back	-0.711	0.328	-2.170	0.191
S3 - S2	DBS	back	0.302	0.257	1.177	0.765
S3 - F2	DBS	back	-0.591	0.272	-2.170	0.191
S3 - F3	DBS	back	-0.744	0.357	-2.085	0.227
S2 - F2	DBS	back	-0.893	0.271	-3.298	0.009
S2 - F3	DBS	back	-1.046	0.356	-2.935	0.028
F2 - F3	DBS	back	-0.153	0.356	-0.429	0.993

Note: These comparisons demonstrate estimated differences between speech rates for each group and vowel backness. Results are averaged over listener. Results are given on the log odds ratio scale. Contrast, group, estimated difference, standard errors, z-ratio, and Tukey-adjusted significances are reported.

Appendix G: Experiment 3

Table 23: Summary of fixed effect coefficients for the sentence intelligibility model.

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
(INTERCEPT)	2.127	0.065	77.190	32.624	0.000
GROUP0	-0.798	0.147	76.481	-5.420	0.000
GROUP1	-1.092	0.160	76.451	-6.820	0.000
GROUP2	-0.469	0.187	76.387	-2.513	0.014
RATE.S3	-0.303	0.052	88.157	-5.810	0.000
RATE.S2	-0.391	0.049	77.817	-7.987	0.000
RATE.F2	0.446	0.065	145.798	6.891	0.000
RATE.F3	0.879	0.233	1358.425	3.776	0.000
NWORDS.1	-0.175	0.043	2405.135	-4.033	0.000
NWORDS.2	-0.179	0.022	2395.302	-8.139	0.000
NWORDS.3	0.066	0.017	2397.727	3.796	0.000
NWORDS.4	-0.035	0.012	2380.836	-2.996	0.003
NWORDS.5	-0.076	0.010	2403.960	-7.787	0.000
GROUP0:RATE.S3	-0.296	0.109	67.807	-2.709	0.009
GROUP1:RATE.S3	-0.019	0.122	74.843	-0.155	0.877
GROUP2:RATE.S3	-0.210	0.147	82.858	-1.427	0.157
GROUP0:RATE.S2	-0.477	0.113	82.966	-4.234	0.000
GROUP1:RATE.S2	-0.101	0.118	71.762	-0.859	0.393
GROUP2:RATE.S2	-0.096	0.138	71.817	-0.698	0.488
GROUP0:RATE.F2	0.265	0.119	70.397	2.230	0.029
GROUP1:RATE.F2	0.335	0.130	71.136	2.576	0.012
GROUP2:RATE.F2	0.267	0.158	75.427	1.693	0.095
GROUP0:RATE.F3	0.541	0.168	81.963	3.211	0.002
GROUP1:RATE.F3	0.491	0.181	78.868	2.716	0.008
GROUP2:RATE.F3	0.555	0.212	79.111	2.623	0.010
RATE.S3:NWORDS.1	0.192	0.074	2397.684	2.604	0.009
RATE.S2:NWORDS.1	0.200	0.070	2426.736	2.840	0.005
RATE.F2:NWORDS.1	-0.049	0.063	2439.174	-0.780	0.435
RATE.F3:NWORDS.1	0.052	0.092	2110.554	0.564	0.573
RATE.S3:NWORDS.2	0.213	0.035	2404.315	5.995	0.000
RATE.S2:NWORDS.2	0.107	0.036	2433.033	2.974	0.003
RATE.F2:NWORDS.2	-0.027	0.045	2421.469	-0.600	0.548

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
RATE.F3:NWORDS.2	-0.326	0.318	2150.863	-1.025	0.305
RATE.S3:NWORDS.3	-0.082	0.030	2391.192	-2.722	0.007
RATE.S2:NWORDS.3	-0.020	0.028	2420.448	-0.724	0.469
RATE.F2:NWORDS.3	-0.031	0.026	2425.158	-1.192	0.233
RATE.F3:NWORDS.3	0.080	0.083	2232.697	0.964	0.335
RATE.S3:NWORDS.4	0.032	0.019	2391.024	1.712	0.087
RATE.S2:NWORDS.4	0.005	0.020	2419.155	0.247	0.805
RATE.F2:NWORDS.4	-0.018	0.021	2421.257	-0.828	0.408
RATE.S3:NWORDS.5	0.084	0.014	2440.142	5.899	0.000
RATE.S2:NWORDS.5	0.022	0.016	2455.548	1.388	0.165
RATE.F2:NWORDS.5	-0.002	0.039	2397.593	-0.052	0.959

Note: GROUP0 YC vs. OC, PD, DBS; GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). NWORDS = Number of words in the sentence, where 1 = 5 vs. 6 words, 2 = 5-6 vs. 7, 3 = 5-7 vs. 8, 4 = 5-8 vs. 9, 5 = 5-9 vs. 10. Coefficient estimates, standard errors, degrees of freedom (df), t-values, and significances

Table 24: Summary of fixed effect coefficients for the conversational intelligibility model.

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
(INTERCEPT)	1.995	0.076	102.333	26.299	0.000
GROUP0	-1.607	0.174	102.333	-9.236	0.000
GROUP1	-1.307	0.185	102.333	-7.056	0.000
GROUP2	-0.793	0.217	102.333	-3.662	0.000
RATES4	0.085	0.077	98.216	1.110	0.270
RATES3	-0.044	0.070	100.492	-0.637	0.526
RATES2	0.030	0.066	89.975	0.460	0.646
RATEF2	0.323	0.067	96.921	4.794	0.000
RATEF3	0.392	0.076	93.901	5.179	0.000
RATEF4	0.626	0.082	84.548	7.643	0.000
GROUP0:RATES4	0.734	0.172	98.837	4.268	0.000
GROUP1:RATES4	0.221	0.185	98.534	1.195	0.235
GROUP2:RATES4	0.404	0.226	97.214	1.791	0.076
GROUP0:RATES3	0.201	0.160	100.492	1.258	0.211
GROUP1:RATES3	0.031	0.170	100.492	0.182	0.856
GROUP2:RATES3	0.457	0.199	100.492	2.297	0.024
GROUP0:RATES2	0.206	0.151	89.975	1.362	0.177
GROUP1:RATES2	0.211	0.161	89.975	1.310	0.193
GROUP2:RATES2	0.511	0.188	89.975	2.719	0.008
GROUP0:RATEF2	0.512	0.154	96.804	3.335	0.001
GROUP1:RATEF2	-0.122	0.164	96.865	-0.747	0.457
GROUP2:RATEF2	0.325	0.194	97.063	1.674	0.097
GROUP0:RATEF3	0.587	0.172	93.948	3.406	0.001
GROUP1:RATEF3	-0.084	0.184	93.925	-0.456	0.650
GROUP2:RATEF3	0.428	0.218	93.818	1.964	0.053
GROUP0:RATEF4	0.824	0.185	84.811	4.460	0.000
GROUP1:RATEF4	-0.048	0.198	84.681	-0.244	0.808
GROUP2:RATEF4	0.534	0.239	84.141	2.232	0.028

Note: GROUP0 YC vs. OC, PD, DBS; GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). Coefficient estimates, standard errors, degrees of freedom (df), t-values, and significances are reported.

Appendix H: Acoustic correlates of intelligibility

Table 25: Summary of fixed effect coefficients for the acoustic model of intelligibility.

Contrast	Estimate	Std. Error	df	t value	Pr(>|t|)
(INTERCEPT)	-0.183	0.062	118.021	-2.973	0.004
GROUP1	0.358	0.111	93.792	3.238	0.002
GROUP2	0.003	0.130	97.432	0.027	0.979
RATE.S3	-0.182	0.057	172.469	-3.218	0.002
RATE.S2	-0.186	0.052	1799.528	-3.609	0.000
RATE.F2	-0.073	0.046	1820.349	-1.578	0.115
RATE.F3	0.003	0.060	98.062	0.044	0.965
VOICING	-0.097	0.034	1758.662	-2.834	0.005
VOT	-1.728	1.215	1704.582	-1.423	0.155
VDC	-0.103	0.081	1563.401	-1.275	0.202
QVAI	0.747	0.167	450.285	4.484	0.000
INT	0.043	0.007	619.801	5.758	0.000
HNR	-0.028	0.004	1544.729	-7.267	0.000
SEX	0.161	0.051	58.063	3.150	0.003
CLIPPING	-0.115	0.114	891.313	-1.009	0.313
GROUP1:RATE.S3	-0.185	0.098	103.907	-1.890	0.062
GROUP2:RATE.S3	-0.143	0.117	106.281	-1.219	0.226
GROUP1:RATE.S2	0.007	0.099	1773.241	0.072	0.943
GROUP2:RATE.S2	-0.044	0.112	1741.291	-0.397	0.692
GROUP1:RATE.F2	-0.005	0.091	1732.752	-0.060	0.952
GROUP2:RATE.F2	-0.206	0.110	1735.998	-1.864	0.062
GROUP1:RATE.F3	-0.076	0.111	70.239	-0.684	0.497
GROUP2:RATE.F3	-0.397	0.139	67.701	-2.864	0.006
VOICING:VOT	-3.829	0.975	1824.402	-3.925	0.000
VDC:VOICING	0.015	0.062	1811.835	0.248	0.804

Note: GROUP1 = OC vs. PD, DBS; GROUP2 = PD vs. DBS. All modified rates (S3, S2, F2, F3) are compared to habitual (H1). Coefficient estimates, standard errors, degrees of freedom (df), z-ratios, and significances are reported.

REFERENCES

Abeyesekera, A., Adams, S., Mancinelli, C., Knowles, T., Gilmore, G., Delrobaei, M., & Jog, M. (2019). Effects of deep brain stimulation of the subthalamic nucleus settings on voice quality, intensity, and prosody in Parkinson's disease: Preliminary evidence for speech optimization. *Canadian Journal of Neurological Sciences / Journal Canadien Des Sciences Neurologiques*, *46*(3), 1–8. doi:10.1017/cjn.2019.16

Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, *63*, 75–86.

Ackermann, H., Hertrich, I., & Hehr, T. (1995). Oral diadochokinesis in neurological dysarthrias. *Folia Phoniatrica et Logopaedica*, *47*(1), 15–23. doi:10.1159/000266338

Ackermann, H., & Ziegler, W. (1991). Articulatory deficits in parkinsonian dysarthria: An acoustic analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*, 1093–1098. doi:10.1136/jnnp.54.12.1093

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech velocity profiles. Journal of Speech and Hearing Research, 36(1), 41–54. doi:10.1044/jshr.3601.41

Adams, S. G. (1994). Accelerating speech in a case of hypokinetic dysarthria: Descriptions and treatment. In J. A. Till, K. M. Yorkston, & D. R. Beukelman (Eds.), *Motor speech disorders: Advances in assessment and treatment* (pp. 213–228). Baltimore: Paul A. Brookes.

Adams, S. G., & Dykstra, A. D. (2009). Hypokinetic dysarthria. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders*. New York: Thieme Publishing Group.

Adams, S. G., Dykstra, A., Jenkins, M., & Jog, M. (2008). Speech-to-noise levels and conversational intelligibility in hypophonia and Parkinson's disease. *Journal of Medical Speech-Language Pathology*, *16*(4), 165–172.

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech, Language, and Hearing Research*, *36*(1), 41–54. doi:10.1044/jshr.3601.41

Adams, S. G., Winnell, J., & Jog, M. (2010). Effects of interlocutor distance, multi-talker background noise, and a concurrent manual task on speech intensity in Parkinson's disease. *Journal of Medical Speech-Language Pathology*, *18*(4), 1–9.

Adams, S., Haralabous, O., Dykstra, A., Abrams, K., & Jog, M. (2005). Effects of multi-talker background noise on the intensity of spoken sentences in Parkinson's disease. *Canadian Acoustics*, *33*(3), 94–95.

Agwuele, A., Sussman, H. M., & Lindblom, B. (2008). The effect of speaking rate on consonant vowel coarticulation. *Phonetica*, *65*(4), 194–209.

Aldridge, D., Theodoros, D., Angwin, A., & Vogel, A. P. (2016). Speech outcomes in Parkinson's disease after subthalamic nucleus deep brain stimulation: A systematic review. *Parkinsonism & Related Disorders*, *33*, 3–11. doi:10.1016/j.parkreldis.2016.09.022

Allison, K. M., Yunusova, Y., & Green, J. R. (2019). Shorter sentence length maximizes intelligibility and speech motor performance in persons with dysarthria due to amyotrophic lateral sclerosis. *American Journal of Speech-Language Pathology*, *28*(1), 96–107.

Altmann, L. J., & Troche, M. S. (2011). High-level language production in Parkinson's disease: A review. *Parkinson's Disease*, *2011*. doi:10.4061/2011/238956

Amano-Kusumoto, A., & Hosom, J.-P. (2011). A review of research on speech intelligibility and correlations with acoustic features. *Center for Spoken Language Understanding, CSLU-011-001*.

Amerman, J. D., & Parnell, M. M. (1992). Speech timing strategies in elderly adults. *Journal of Phonetics*.

Antolík, T., & Fougeron, C. (2013). Consonant distortions in dysarthria due to Parkinson's disease, amyotrophic lateral sclerosis and cerebellar ataxia. In *Interspeech 2013*.

Aquino, C. C., & Fox, S. H. (2015). Clinical spectrum of levodopa-induced complications. *Movement Disorders*, *30*(1), 80–89. doi:10.1002/mds.26125

Artieda, J., Pastor, M. A., Lacruz, F., & Obeso, J. A. (1992). Temporal discrimination is abnormal in Parkinson's disease. *Brain*, *115*(1), 199–210. doi:10.1093/brain/115.1.199

Atkinson-Clement, C., Sadat, J., & Pinto, S. (2015). Behavioral treatments for speech in Parkinson's disease: Meta-analyses and review of the literature. *Neurodegenerative Disease Management*, *5*(3), 233–248.

Auzou, P., Ozsancak, C., Morris, R. J., Jan, M., Eustache, F., & Hannequin, D. (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: A review. *Clinical Linguistics & Phonetics*, *14*(2), 131–150.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31–56.

Åkesson, J., Lindh, J., & Hartelius, L. (2010). Post surgery effects on VOT for Parkinson disease STN/DBS patients. *Proceedings from FONETIK 2010, Working Papers*, *54*, 119–124.

Baker, R., & Hazan, V. (2011). DiapixUK: Task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, *43*(3), 761–770.

Bang, Y.-I., Min, K., Sohn, Y. H., & Cho, S.-R. (2013). Acoustic characteristics of vowel sounds in patients with Parkinson disease. *NeuroRehabilitation*, *32*(3), 649–654.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*(1), 1–22.

Bard, E. G., & Aylett, M. P. (1999). The dissociation of deaccenting, givenness, and syntactic role in spontaneous speech. In *Proceedings of the XIVth international congress of phonetic sciences* (Vol. 3, pp. 1753–6). University of California Berkeley, CA.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92–111.

Benjamin, B. J. (1982). Phonological performance in gerontological speech. *Journal of Psycholinguistic Research*, *11*(2), 159–167.

Berg, D., Postuma, R. B., Adler, C. H., Bloem, B. R., Chan, P., Dubois, B., … others. (2015). MDS research criteria for prodromal Parkinson's disease. *Movement Disorders*, *30*(12), 1600–1611.

Beukelman, D., Yorkston, K., & Tice, R. (1997). Pacer/tally rate measurement software. *Lincoln, NE: Tice Technology Services*.

Beverly, D., Cannito, M. P., Chorna, L., Wolf, T., Suiter, D. M., & Bene, E. R. (2010). Influence of stimulus sentence characteristics on speech intelligibility scores in hypokinetic dysarthria. *Journal of Medical Speech-Language Pathology*, *18*(4), 9–14.

Blanchet, P. G., & Snyder, G. J. (2009). Speech rate deficits in individuals with Parkinson's disease: A review of the literature. *Journal of Medical Speech-Language Pathology*, *17*(1), 1–8.

Blanchet, P. G., & Snyder, G. J. (2010). Speech rate treatments for individuals with dysarthria: a tutorial. *Percept Mot Skills*, *110*(3 Pt 1), 965–982. doi:10.2466/pms.110.3.965-982

Boberg, C. (2005). The Canadian shift in Montreal. *Language Variation and Change*, *17*(2), 133–154.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer [computer program] (Version 5.3). http://www.praat.org/.

Bonnet, A.-M., Loria, Y., Saint-Hilaire, M.-H., Lhermitte, F., & Agid, Y. (1987). Does long-term aggravation of Parkinson's disease result from nondopaminergic lesions? *Neurology*, *37*(9), 1539–1539.

Braak, H., Ghebremedhin, E., Rüb, U., Bratzke, H., & Del Tredici, K. (2004). Stages in the development of Parkinson's disease-related pathology. *Cell and Tissue Research*, *318*(1), 121–134. doi:10.1007/s00441-004-0956-9

Bradlow, A., Nygaard, L., & Pisony, D. (1995). On the contribution of instance-specific characteristics to speech perception. In C. Sorin, J. Mariani, H. Meloni, & J. Schoentgen (Eds.), *Levels in speech communication: Relations and interactions* (pp. 13–25). Amsterdam: Elsevier.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, *20*(3), 255–272.

Breitenstein, C., Van Lancker, D., Daum, I., & Waters, C. H. (2001). Impaired perception of vocal emotions in Parkinson's disease: Influence of speech time processing and executive functioning. *Brain and Cognition*, *45*(2), 277–314.

Bryan, J., Luszcz, M. A., & Crawford, J. R. (1997). Verbal knowledge and speed of information processing as mediators of age differences in verbal fluency performance among older adults. *Psychology and Aging*, *12*(3), 473. doi:10.1037//0882-7974.12.3.473

Buccheri, R. A. (2013). *Effects of speaking mode (clear, habitual, slow speech) on vowels and intelligibility of individuals with Parkinson's disease* (PhD thesis). The City University of New York.

Bunton, K. (1999). *An acoustic item-analysis of speech intelligibility* (PhD thesis). University of Wisconsin-Madison.

Bunton, K. (2006). Fundamental frequency as a perceptual cue for vowel identification in speakers with Parkinson's disease. *Folia Phoniatrica et Logopaedica*, *58*(5), 323–339. doi:10.1159/000094567

Bunton, K. (2008). Speech versus nonspeech: Different tasks, different neural organization. *Seminars in Speech and Language*, *29*(4), 267–275. doi:10.1055/s-0028-1103390

Bunton, K., & Keintz, C. K. (2008). The use of a dual-task paradigm for assessing speech intelligibility in clients with Parkinson disease. *Journal of Medical Speech-Language Pathology*, *16*(3), 141.

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, and Hearing Research*, *50*(6), 1481–1495. doi:10.1044/1092-4388(2007/102)

Bunton, K., & Weismer, G. (2001). The relationship between perception and acoustics for a high-low vowel contrast produced by speakers with dysarthria. *Journal of Speech, Language, and Hearing Research*, *44*(6), 1215–1228. doi:10.1044/1092-4388(2001/095)

Bunton, K., & Weismer, G. (2002). Segmental level analysis of laryngeal function in persons with motor speech disorders. *Folia Phoniatrica et Logopaedica*, *54*(5), 223–239.

Burke, D. M., & MacKay, D. G. (1997). Memory, language, and ageing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1363), 1845–1856.

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, *15*(1-2), 39–54.

Caligiuri, M. P. (1989). The influence of speaking rate on articulatory hypokinesia in Parkinsonian dysarthria. *Brain and Language*, *36*(3), 493–502. doi:10.1016/0093-934x(89)90080-1

Cannito, M. P., Suiter, D. M., Beverly, D., Chorna, L., Wolf, T., & Pfeiffer, R. M. (2012). Sentence intelligibility before and after voice treatment in speakers with idiopathic Parkinson's disease. *Journal of Voice*, *26*(2), 214–219.

Canter, G. J. (1963). Speech characteristics of patients with Parkinson's disease: I. Intensity, pitch, and duration. *Journal of Speech & Hearing Disorders*.

Canter, G. J. (1965). Speech characteristics of patients with Parkinson's disease: III. Articulation, diadochokinesis, and over-all speech adequacy. *Journal of Speech and Hearing Disorders*, *30*(3), 217–224.

Carter, C. R., Yorkston, K. M., Strand, E. A., & Hammen, V. L. (1996). Effects of semantic and syntactic context on actual and estimated sentence intelligibility of dysarthric speakers. In D. A. Robin, K. M. Yorkston, & D. R. Beukelman (Eds.), *Disorders of motor speech: Assessment, treatment, and clinical characterization*. Paul H Brookes Publishing Company.

Cheesman, M. F., & Jamieson, D. G. (1996). Development, evaluation and scoring of a nonsense word test suitable for use with speakers of Canadian English. *Canadian Acoustics*, *24*(1), 3–11.

Chenausky, K., MacAuslan, J., & Goldhor, R. (2011). Acoustic analysis of PD speech. *Parkinson's Disease*, *2011*, 1–13. doi:10.4061/2011/435232

Cho, T., Whalen, D. H., & Docherty, G. (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics*, *72*, 52–65. doi:10.1016/j.wocn.2018.11.002

Clark, J. P., Adams, S. G., Dykstra, A. D., Moodie, S., & Jog, M. (2014). Loudness perception and speech intensity control in Parkinson's disease. *Journal of Communication Disorders*, *51*, 1–12. doi:10.1016/j.jcomdis.2014.08.001

Clarke, S., Elms, F., & Youssef, A. (1995). The third dialect of English: Some Canadian evidence. *Language Variation and Change*, *7*(2), 209–228.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Connolly, B. S., & Lang, A. E. (2014). Pharmacological treatment of Parkinson disease: A review. *JAMA*, *311*(16), 1670–1683. doi:10.1001/jama.2014.3654

Connor, N. P., Abbs, J. H., Cole, K. J., & Gracco, V. L. (1989). Parkinsonian deficits in serial multiarticulate movements for speech. *Brain*, *112*(4), 997–1009. doi:10.1093/brain/112.4.997

Cox, R. M., Alexander, G. C., & Gilmore, C. (1987). Intelligibility of average talkers in typical listening environments. *The Journal of the Acoustical Society of America*, *81*(5), 1598–1608.

Cushnie-Sparrow, D., Adams, S., Abeyesekera, A., Pieterman, M., Gilmore, G., & Jog, M. (2018). Voice quality severity and responsiveness to levodopa in Parkinson's disease. *Journal of Communication Disorders*, *76*, 1–10. doi:10.1016/j.jcomdis.2018.07.003

Cushnie-Sparrow, D., Adams, S., Knowles, T., Leszcz, T. M., & Jog, M. (2016). Effects of multi-talker noise on the acoustics of voiceless stop consonants in Parkinson's disease. *Western Papers in Linguistics/Cahiers Linguistiques de Western*, *3*(1).

Dagenais, P. A., Brown, G. R., & Moore, R. E. (2006). Speech rate effects upon intelligibility and acceptability of dysarthric speech. *Clinical Linguistics & Phonetics*, *20*(2-3), 141–148.

D'Alatri, L., Paludetti, G., Contarino, M. F., Galla, S., Marchese, M. R., & Bentivoglio, A. R. (2008). Effects of bilateral subthalamic nucleus stimulation and medication on parkinsonian speech impairment. *Journal of Voice*, *22*(3), 365–372. doi:10.1016/j.jvoice.2006.10.010

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech, Language, and Hearing Research*, *12*(3), 462–496. doi:doi:10.1044/jshr.1203.462

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Differential diagnostic patterns of dysarthria. *Journal of Speech, Language, and Hearing Research*, *12*(2), 246. doi:10.1044/jshr.1202.246

Darling, M., & Huber, J. E. (2011). Changes to articulatory kinematics in response to loudness cues in individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *54*(5), 1247–1259. doi:10.1044/1092-4388(2011/10-0024)

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, *54*, 35–50. doi:10.1016/j.wocn.2015.09.003

Davidson, L. (2018). Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association*, *48*(3), 331–356.

De Bodt, M. S., Huici, M. E. H.-D., & Van De Heyning, P. H. (2002). Intelligibility as a linear combination of dimensions in dysarthric speech. *Journal of Communication Disorders*, *35*(3), 283–292. doi:10.1016/s0021-9924(02)00065-5

De Lau, L. M., & Breteler, M. M. (2006). Epidemiology of Parkinson's disease. *The Lancet Neurology*, *5*(6), 525–535. doi:10.1016/s1474-4422(06)70471-9

DePaul, R., & Kent, R. D. (2000). A longitudinal case study of ALS: Effects of listener familiarity and proficiency on intelligibility judgments. *American Journal of Speech-Language Pathology*, *9*(3), 230–240.

Deuschl, G., Schade-Brittinger, C., Krack, P., Volkmann, J., Schäfer, H., Bötzel, K., … others. (2006). A randomized trial of deep-brain stimulation for Parkinson's disease. *New England Journal of Medicine*, *355*(9), 896–908. doi:10.1056/nejmoa060281

D'Innocenzo, J., Tjaden, K., & Greenman, G. (2006). Intelligibility in dysarthria: Effects of listener familiarity and speaking condition. *Clinical Linguistics & Phonetics*, *20*(9), 659–675.

Downie, A., Low, J., & Lindsay, D. (1981). Speech disorder in parkinsonism: Usefulness of delayed auditory feedback in selected cases. *International Journal of Language & Communication Disorders*, *16*(2), 135–139.

Drager, K. D., & Reichle, J. E. (2001). Effects of discourse context on the intelligibility of synthesized speech for young adult and older adult listeners. *Journal of Speech, Language, and Hearing Research*.

Dromey, C., & Bjarnason, S. (2011). A preliminary report on disordered speech with deep brain stimulation in individuals with Parkinson's disease. *Parkinson's Disease*, *2011*, 1–11. doi:10.4061/2011/796205

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate their impact on measures of respiration, phonation, and articulation. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1003–1018.

Dromey, C., Ramig, L. O., & Johnson, A. B. (1995). Phonatory and articulatory changes associated with increased vocal intensity in patients with Parkinson disease. *Journal of Speech and Hearing Research*, *38*, 751–764. doi:10.1044/jshr.3804.751

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). Elsevier Health Sciences.

Dumer, A. I., Oster, H., McCabe, D., Rabin, L. A., Spielman, J. L., Ramig, L. O., & Borod, J. C. (2014). Effects of the Lee Silverman Voice Treatment (LSVT LOUD) on hypomimia in Parkinson's disease. *Journal of the International Neuropsychological Society*, *20*(3), 302–312. doi:10.1017/s1355617714000046

Dworkin, J. P., & Aronson, A. E. (1986). Tongue strength and alternate motion rates in normal and dysarthric subjects. *Journal of Communication Disorders*, *19*(2), 115–132. doi:10.1016/0021-9924(86)90015-8

Dykstra, A. D., Adams, S. G., & Jog, M. (2015). Examining the relationship between speech intensity and self-rated communicative effectiveness in individuals with Parkinson's disease and hypophonia. *Journal of Communication Disorders*, *56*, 103–112. doi:10.1016/j.jcomdis.2015.06.012

Eklund, E., Qvist, J., Sandström, L., Viklund, F., van Doorn, J., & Karlsson, F. (2014). Perceived articulatory precision in patients with Parkinson's disease after deep brain stimulation of subthalamic nucleus and caudal zona incerta. *Clinical Linguistics & Phonetics*, *29*(2), 150–166.

Elgamal, S. A., Roy, E. A., & Sharratt, M. T. (2011). Age and verbal fluency: The mediating effect of speed of processing. *Canadian Geriatrics Journal*, *14*(3), 66. doi:10.5770/cgj.v14i3.17

El Sharkawi, A. (2002). Swallowing and voice effects of Lee Silverman Voice Treatment (LSVT): A pilot study. *Journal of Neurology, Neurosurgery & Psychiatry*, *72*(1), 31–36. doi:10.1136/jnnp.72.1.31

Engstrand, O. (1988). Articulatory correlates of stress and speaking rate in Swedish VCV utterances. *The Journal of the Acoustical Society of America*, *83*(5), 1863–1875.

Enos, N. M., Abur, D., & Stepp, C. E. (2018). *Comparing visual-analog-scale ratings and orthographic transcription estimates of intelligibility in naive listeners*.

Fabbri, M., Guimarães, I., Cardoso, R., Coelho, M., Guedes, L. C., Rosa, M. M., … others. (2017). Speech and voice response to a levodopa challenge in late-stage Parkinson's disease. *Frontiers in Neurology*, *8*, 432.

Fahn, S. (2003). Description of Parkinson's disease as a clinical syndrome. *Annals of the New York Academy of Sciences*, *991*(1), 1–14.

Fant, G. (1966). A note on vocal tract size factors and non-uniform f-pattern scalings. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *1*, 22–30.

Fant, G. (1970). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations*. Walter de Gruyter.

Fant, G. (1975). Non-uniform vowel normalization. *STL-QPSR*, *16*(2-3), 1–19.

Farris, S., & Giroux, M. (2013). Retrospective review of factors leading to dissatisfaction with subthalamic nucleus deep brain stimulation during long-term management. *Surgical Neurology International*. doi:10.4103/2152-7806.112612

Feenaughty, L., Tjaden, K., & Sussman, J. (2014). Relationship between acoustic measures and judgments of intelligibility in Parkinson's disease: A within-speaker approach. *Clinical Linguistics & Phonetics*, *28*(11), 857–878.

Fenoy, A. J., McHenry, M. A., & Schiess, M. C. (2016). Speech changes induced by deep brain stimulation of the subthalamic nucleus in Parkinson disease: Involvement of the dentatorubrothalamic tract. *Journal of Neurosurgery*, 1–11.

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *112*(1), 259–271.

Ferguson, S. H., & Quené, H. (2014). Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *135*(6), 3570–3584.

Ferrand, C. T. (2002). Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice*, *16*(4), 480–487.

Fischer, E., & Goberman, A. M. (2010). Voice onset time in Parkinson disease. *Journal of Communication Disorders*, *43*(1), 21–34. doi:10.1016/j.jcomdis.2009.07.004

Fisher, H. B., & Logemann, J. A. (1971). *The Fisher-Logemann Test of Articulation Competence*. Pro-Ed.

Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., & Liss, J. M. (2015). The relationship between speech segment duration and vowel centralization in a group of older speakers. *The Journal of the Acoustical Society of America*, *138*(4), 2132–2139. doi:10.1121/1.4930563

Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., & Liss, J. M. (2017a). Assessing vowel centralization in dysarthria: A comparison of methods. *Journal of Speech, Language, and Hearing Research*, 1–14. doi:10.1044/2016_JSLHR-S-15-0355

Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., Sinex, D. G., & Liss, J. M. (2017b). Predicting intelligibility gains in individuals with dysarthria from baseline speech features. *Journal of Speech, Language, and Hearing Research*, *60*(11), 1–15. doi:10.1044/2016_jslhr-s-16-0218

Fletcher, H. (1953). *Speech and hearing in communication*. D. van Nostrand.

Flint, A. J., Black, S. E., Campbell-Taylor, I., Gailey, G. F., & Levinton, C. (1992). Acoustic analysis in the differentiation of Parkinson's disease and major depression. *Journal of Psycholinguistic Research*, *21*(5), 383–399.

Fogerty, D., & Kewley-Port, D. (2009). Perceptual contributions of the consonant-vowel boundary to sentence intelligibility. *The Journal of the Acoustical Society of America*, *126*(2), 847–857. doi:10.1121/1.2972133

Forrest, K., Nygaard, L., Pisoni, D. B., & Siemers, E. (1998). Effects of speaking rate on word recognition in Parkinson's disease and normal aging. *Journal of Medical Speech-Language Pathology*, *6*(1), 1.

Forrest, K., Weismer, G., & Turner, G. S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults. *The Journal of the Acoustical Society of America*, *85*(6), 2608–2622. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2745883

Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical Society of America*, *90*(4), 1816–1827. doi:10.1121/1.401662

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, *8*(1), 113–133.

Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech*, *31*(4), 307–319.

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, *36*(2-3), 171–195.

Fox, C. M., & Ramig, L. O. (1997). Vocal sound pressure level and self-perception of speech and voice in men and women with idiopathic Parkinson's disease. *American Journal of Speech-Language Pathology*, *6*(2), 85–94. doi:10.1044/1058-0360.0602.85

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (Second.). Thousand Oaks CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Gamboa, J., Jiménez-Jiménez, F. J., Nieto, A., Montojo, J., Ortí-Pareja, M., Molina, J. A., … Cobeta, I. (1997). Acoustic voice analysis in patients with Parkinson's disease treated with dopaminergic drugs. *Journal of Voice*, *11*(3), 314–320. doi:10.1016/s0892-1997(97)80010-0

Gay, T. (1978). Effect of speaking rate on vowel formant movements. *The Journal of the Acoustical Society of America*, *63*(1), 223–230.

Gentil, M., Chauvin, P., Pinto, S., Pollak, P., & Benabid, A.-L. (2001). Effect of bilateral stimulation of the subthalamic nucleus on parkinsonian voice. *Brain and Language*, *78*, 233–240. Retrieved from http://www.sciencedirect.com/science/article/pii/S0093934X01924666

Gentil, M., Pinto, S., Pollak, P., & Benabid, A.-L. (2003). Effect of bilateral stimulation of the subthalamic nucleus on parkinsonian dysarthria. *Brain and Language*, *85*(2), 190–196. doi:10.1016/S0093-934X(02)00590-4

Gibb, W., & Lees, A. (1988). The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *51*(6), 745–752.

Glosser, G., & Deser, T. (1992). A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *Journal of Gerontology*, *47*(4), P266–P272.

Gonera, E. G., Hof, M. V., Berger, H. J., Weel, C. van, & Horstink, M. W. (1997). Symptoms and duration of the prodromal phase in Parkinson's disease. *Movement Disorders*, *12*(6), 871–876. doi:10.1002/mds.870120607

Goozee, J. V., Stephenson, D. K., Murdoch, B. E., Darnell, R. E., & Lapointe, L. L. (2005). Lingual kinematic strategies used to increase speech rate: Comparison between younger and older adults. *Clinical Linguistics & Phonetics*, *19*, 319–334.

Gordon-Salant, S. (1986). Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. *The Journal of the Acoustical Society of America*, *80*(6), 1599–1607.

Graças, R. R. D., Gama, A. C. C., Cardoso, F. E. C., Lopes, B. P., & Bassi, I. B. (2012). Objective and subjective analysis of women's voice with idiopathic Parkinson's disease. *Arquivos de Neuro-Psiquiatria*.

Grant, M. J., Button, C. M., & Snook, B. (2017). An evaluation of interrater reliability measures on binary tasks using d-Prime. *Applied Psychological Measurement*, *41*(4), 264–276.

Gutek, J., Rochet, A. P., Robin, D., Yorkston, K., & Beukelman, D. (1996). Effects of insertion of interword pause on the intelligibility of dysarthric speech. *Disorders of Motor Speech: Assessment, Treatment, and Clinical Characterization*, 105–120.

Hall, Z. D. (2013). *Effect of rate reduction on speech intelligibility in individuals with dysarthria* (Master's thesis). Louisiana State University; Agricultural; Mechanical College.

Hammen, V. L., & Yorkston, K. M. (1996). Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria. *Journal of Communication Disorders*, *29*(6), 429–445. doi:10.1016/0021-9924(95)00037-2

Hammen, V. L., Yorkston, K. M., & Minifie, F. D. (1994). Effects of temporal alterations on speech intelligibility in parkinsonian dysarthria. *Journal of Speech, Language, and Hearing Research*, *37*(2), 244–253. doi:10.1044/jshr.3702.244

Hammen, V., Yorkston, K., & Dowden, P. (1991). Index of contextual intelligibility: Impact of semantic context in dysarthria. *Dysarthria and Apraxia of Speech: Perspectives on Management*, 43–53.

Hanson, W. R., & Metter, E. J. (1983). DAF speech rate modification in Parkinson's disease: A report of two cases. In *Clinical dysarthria* (pp. 231–251). College-Hill, San Diego, California.

Harel, D., & McAllister, T. (2019). Multilevel models for communication sciences and disorders. *Journal of Speech, Language, and Hearing Research*, 1–19.

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, *36*(4), 664–679.

Hegland, K. W., Troche, M., & Brandimore, A. (2019). Relationship between respiratory sensory perception, speech, and swallow in Parkinson's disease. *Movement Disorders Clinical Practice*, *6*(3), 243–249.

Henton, C. (1995). Cross-language variation in the vowels of female and male speakers. In *Proceedings of the XIIIth international congress of phonetic sciences* (Vol. 4, pp. 420–423).

Hertrich, I., & Ackermann, H. (1995a). Coarticulation in slow speech: Durational and spectral analysis. *Language and Speech*, *38*(2), 159–187.

Hertrich, I., & Ackermann, H. (1995b). Gender-specific vocal dysfunctions in Parkinson's disease: Electroglottographic and acoustic analyses. *Annals of Otology, Rhinology & Laryngology*, *104*(3), 197–202.

Higgins, M. B., Netsell, R., & Schulte, L. (1998). Vowel-related differences in laryngeal articulatory and phonatory function. *Journal of Speech, Language, and Hearing Research*, *41*(4), 712–724. doi:10.1044/jslhr.4104.712

Hirata, Y., & Tsukada, K. (2004). The effects of speaking rates and vowel length on formant movements in Japanese. In *Proceedings of the 2003 Texas Linguistics Society Conference* (pp. 73–85). Citeseer.

Hirose, H., Kiritani, S., & Sawashima, M. (1982). Velocity of articulatory movements in normal and dysarthric subjects. *Folia Phoniatrica et Logopaedica*, *34*(4), 210–215. doi:10.1159/000265651

Ho, A. K., Bradshaw, J. L., Iansek, R., & Alfredson, R. (1999). Speech volume regulation in Parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia*, *37*, 1453–1460.

Ho, A. K., Iansek, R., & Bradshaw, J. L. (2001). Motor instability in parkinsonian speech intensity. *Neuropsychiatry, Neurophsychology, and Behavioral Neurology*, *14*(2), 109–116.

Ho, A. K., Iansek, R., & Bradshaw, J. L. (2002). The effect of a concurrent task on Parkinsonian speech. *Journal of Clinical and Experimental Neuropsychology*, *24*(1), 36–47.

Holmes, R. J., Oates, J. M., Phyland, D. J., & Hughes, A. J. (2000). Voice characteristics in the progression of Parkinson's disease. *International Journal of Language & Communication Disorders*, *35*(3), 407–418. doi:10.1080/136828200410654

Honda, K. (1983). Relationship between pitch control and vowel articulation. *Haskins Laboratories Status Report on Speech Research, SR*, *73*, 269–82.

Hsu, S.-C., Jiao, Y., McAuliffe, M. J., Berisha, V., Wu, R.-M., & Levy, E. S. (2017). Acoustic and perceptual speech characteristics of native Mandarin speakers with Parkinson's disease. *The Journal of the Acoustical Society of America*, *141*(3), EL293–EL299.

Huber, J. E. (2008). Effects of utterance length and vocal loudness on speech breathing in older adults. *Respiratory Physiology & Neurobiology*, *164*(3), 323–330. doi:10.1016/j.resp.2008.08.007

Hughes, A. J., Daniel, S. E., Blankson, S., & Lees, A. J. (1992). A clinicopathologic study of 100 cases of Parkinson's disease. *Archives of Neurology*, *50*(2), 140–148.

Hustad, K. C. (2006). Estimating the intelligibility of speakers with dysarthria. *Folia Phoniatrica et Logopaedica*, *58*(3), 217–228.

Hustad, K. C. (2007). Effects of speech stimuli and dysarthria severity on intelligibility scores and listener confidence ratings for speakers with cerebral palsy. *Folia Phoniatrica et Logopaedica*, *59*(6), 306–317.

Hustad, K. C., & Beukelman, D. R. (2001). Effects of linguistic cues and stimulus cohesion on intelligibility of severely dysarthric speech. *Journal of Speech, Language, and Hearing Research*.

Hustad, K. C., & Cahill, M. A. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology*.

Hustad, K. C., & Lee, J. (2008). Changes in speech production associated with alphabet supplementation. *Journal of Speech, Language, and Hearing Research*, *51*, 1438–1450. doi:10.1044/1092-4388(2008/07-0185)

Iddon, H., Read, J., & Miller, N. (2015). Does voice amplification increase intelligibility in people with Parkinson's disease? *International Journal of Therapy and Rehabilitation*, *22*(10), 479–486. doi:10.12968/ijtr.2015.22.10.479

Im, H., Adams, S., Abeyesekera, A., Pieterman, M., Gilmore, G., & Jog, M. (2018). Effect of levodopa on speech dysfluency in Parkinson’s disease. *Movement Disorders Clinical Practice*, *6*(2), 150–154.

Iulianella, I., Adams, S. G., & Gow, A. K. (2008). Effects of sub-thalamic deep brain stimulation on speech production in Parkinson’s disease: A critical review of the literature. *Canadian Journal of Speech-Language Pathology & Audiology*, *32*(2), 85–91.

Jacewicz, E., Fox, R. A., O’Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change*, *21*(02), 233–256.

Jankovic, J. (2008). Parkinson’s disease: Clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*(4), 368–376.

Jiménez-Jiménez, F., & Molina, J. (1997). Pharmacological therapy of complicated Parkinson’s disease. *Revista de Neurologia*, *25*, S170–9.

Karlsson, F., Olofsson, K., Blomstedt, P., Linder, J., Nordh, E., & van Doorn, J. (2014). Articulatory closure proficiency in patients with Parkinson’s disease following deep brain stimulation of the subthalamic nucleus and caudal zona incerta. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1178–1190. doi:10.1044/2014_JSLHR-S-13-0010

Karlsson, F., Unger, E., Wahlgren, S., Blomstedt, P., Linder, J., Nordh, E., … Doorn, J. van. (2011). Deep brain stimulation of caudal zona incerta and subthalamic nucleus in patients with Parkinson’s disease: Effects on diadochokinetic rate. *Parkinson’s Disease*, *2011*.

Karlsson, F., Unger, E., Wahlgren, S., & Doom, J. van. (2012). Treatment effects in voice onset time of plosives associated with deep brain stimulation of the subthalamic nucleus and the caudal zona incerta. *Journal of Medical Speech-Language Pathology*, *20*(4), 65–70.

Karlsson, F., Unger, E., Wahlgren, S., & Doorn, J. van. (2011). Sources of missing data in VOT measurements of patients with Parkinson’s disease under deep brain stimulation in subthalamic nucleus and caudal zona incerta. In *Poster presented at 6th international conference on speech motor control*.

Karlsson, F., & van Doorn, J. (2012). Vowel formant dispersion as a measure of articulation proficiency. *The Journal of the Acoustical Society of America*, *132*(4), 2633–2641. doi:10.1121/1.4746025

Kearney, E., Haworth, B., Scholl, J., Faloutsos, P., Baljko, M., & Yunusova, Y. (2018). Treating speech movement hypokinesia in Parkinson's disease: Does movement size matter? *Journal of Speech, Language, and Hearing Research*, *61*(11), 2703–2721.

Kempler, D., & Van Lancker, D. (2002). Effect of speech task on intelligibility in dysarthria: A case study of Parkinson's disease. *Brain and Language*, *80*(3), 449–464.

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, *5*(3), 7–23.

Kent, R. D., & Rosenbek, J. C. (1982). Prosodic disturbance and neurologic lesion. *Brain and Language*, *15*(2), 259–291.

Kent, R. D., Vorperian, H., Kent, J., & Duffy, J. R. (2003). Voice dysfunction in dysarthria: Application of the Multi-Dimensional Voice Program. *Journal of Communication Disorders*, *36*(4), 281–306. doi:10.1016/s0021-9924(03)00016-9

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *The Journal of Speech and Hearing Disorders*, *54*(4), 482–499. doi:10.1044/jshd.5404.482

Kent, R. D., Weismer, G., Kent, J. F., Vorperian, H. K., & Duffy, J. R. (1999). Acoustic studies of dysarthric speech: Methods, progress, and potential. *Journal of Communication Disorders*, *32*(3), 141–186. doi:10.1016/s0021-9924(99)00004-0

Keshet, J., Sonderegger, M., & Knowles, T. (2014). AutoVOT: A tool for automatic measurement of voice onset time using discriminative structured prediction [Computer program]. Version 0.93. Retrieved from https://github.com/mlml/autovot

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, *25*(2), 143–168. doi:10.1006/jpho.1996.0039

Kim, H., Hasegawa-Johnson, M., & Perlman, A. (2011). Vowel contrast and speech intelligibility in dysarthria. *Folia Phoniatrica et Logopaedica*, *63*(4), 187–194.

Kim, Y., & Choi, Y. (2017). A cross-language study of acoustic predictors of speech intelligibility in individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, 1–13.

Kim, Y., Kent, R. D., & Weismer, G. (2011). An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journal of Speech, Language, and Hearing Research*, *54*(2), 417–429. doi:10.1044/1092-4388(2010/10-0020)

Kim, Y., & Kuo, C. (2012). Effect of level of presentation to listeners on scaled speech intelligibility of speakers with dysarthria. *Folia Phoniatrica et Logopaedica*, *64*(1), 26–33.

Kim, Y., Weismer, G., Kent, R. D., & Duffy, J. R. (2009). Statistical models of F2 slope in relation to severity of dysarthria. *Folia Phoniatrica et Logopaedica*, *61*(6), 329–335. doi:10.1159/000252849

Klawans, H. L. (1986). Individual manifestations of Parkinson's disease after ten or more years of levodopa. *Movement Disorders*, *1*(3), 187–192. doi:10.1002/mds.870010304

Kleinow, J., Smith, A., & Ramig, L. O. (2001). Speech motor stability in IPD effects of rate and loudness manipulations. *Journal of Speech, Language, and Hearing Research*, *44*(5), 1041–1051.

Klostermann, F., Ehlen, F., Vesper, J., Nubel, K., Gross, M., & Marzinzik, F. et al. (2008). Effects of subthalamic deep brain stimulation on dysarthrophonia in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, *79*, 522–529.

Knowles, T., Adams, S., Abeyesekera, A., Mancinelli, C., Gilmore, G., & Jog, M. (2018). Deep brain stimulation of the subthalamic nucleus parameter optimization for vowel acoustics and speech intelligibility in Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *61*(3), 510–524.

Kompoliti, K., Wang, Q., Goetz, C., Leurgans, S., & Raman, R. (2000). Effects of central dopaminergic stimulation by apomorphine on speech in Parkinson's disease. *Neurology*, *54*(2), 458–458.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.

Krack, P., Batir, A., Van Blercom, N., Chabardes, S., Fraix, V., & Ardouin, et a., C. (2003). Five-year follow-up of bilateral stimulation of the subthalamic nucleus in advanced Parkinson's disease. *New England Journal of Medicine*, *349*, 1925–1934. doi:10.1056/nejmoa035275

Krause, J. C., & Braida, L. D. (2002). Investigating alternative forms of clear speech: the effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America*, *112*(5 Pt 1), 2165–2172. doi:10.1121/1.1509432

Kuo, C., Tjaden, K., & Sussman, J. E. (2014). Acoustic and perceptual correlates of faster-than-habitual speech produced by speakers with Parkinson's disease and multiple sclerosis. *Journal of Communication Disorders*, *52*, 156–169. doi:10.1016/j.jcomdis.2014.09.002

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. doi:10.18637/jss.v082.i13

Labov, W., Ash, S., & Boberg, C. (2008). *The atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter.

Lam, J., & Tjaden, K. (2016). Clear speech variants: An acoustic study in Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *59*, 631–646. doi:10.1044/2015_jslhr-s-15-0216

Lam, T. Q., & Watson, D. G. (2010). Repetition is easy: Why repeated referents have reduced prominence. *Memory & Cognition*, *38*(8), 1137–1146.

Lamar, M., Resnick, S. M., & Zonderman, A. B. (2003). Longitudinal changes in verbal memory in older adults: Distinguishing the effects of age from repeat testing. *Neurology*, *60*(1), 82–86.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174.

Lansford, K. L., & Liss, J. M. (2014). Vowel acoustics in dysarthria: Mapping to perception. *Journal of Speech, Language, and Hearing Research*, *57*(1), 68–80. doi:10.1044/1092-4388(2013/12-0263)

Lawson, G. D., & Peterson, M. E. (2011). *Speech audiometry*. Plural Publishing.

LeDorze, G., Dionne, L., Ryalls, J., Julien, M., & Ouellet, L. (1992). The effects of speech and language therapy for a case of dysarthria associated with Parkinson's disease. *International Journal of Language & Communication Disorders*, *27*(4), 313–324.

Lee, J., Dickey, E., & Simmons, Z. (2019). Vowel-specific intelligibility and acoustic patterns in individuals with dysarthria secondary to amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, *62*(1), 34–59. doi:10.1044/2018_jslhr-s-17-0357

Lenth, R. (2018). *Emmeans: Estimated marginal means, aka least-squares means*. Retrieved from https://CRAN.R-project.org/package=emmeans

Lieberman, P., Kako, E., Friedman, J., Tajchman, G., Feldman, L. S., & Jiminez, E. B. (1992). Speech production, syntax comprehension, and cognitive deficits in Parkinson's disease. *Brain and Language*, *43*(2), 169–189. doi:10.1016/0093-934X(92)90127-Z

Limousin, P., Krack, P., & Pollak, P. (1998). Electrical stimulation of the subthalamic nucleus in advanced Parkinson's disease. *New England Journal of Medicine*, *339*, 1105–1111.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America*, *35*(11), 1773–1781.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling* (pp. 403–439). Springer.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422.

Liss, J. M. (1990). Muscle spindles in the human levator veli palatini and palatoglossus muscles. *Journal of Speech, Language, and Hearing Research*, *33*(4), 736–746.

Liss, J. M., Spitzer, S., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *Journal of the Acoustical Society of America*, *112*, 3022–3030. doi:10.1121/1.1515793

Little, M. A., McSharry, P. E., Hunter, E. J., Spielman, J., & Ramig, L. O. (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *IEEE Transactions on Bio-Medical Engineering*, *56*(4), 1015.

Liu, H.-M., Tseng, C.-H., & Tsao, F.-M. (2000). Perceptual and acoustic analysis of speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *Clinical Linguistics & Phonetics*, *14*(6), 447–464.

Logan, K. J., Roberts, R. R., Pretto, A. P., & Morey, M. J. (2002). Speaking slowly: Effects of four self-guided training approaches on adults' speech rate and naturalness. *American Journal of Speech-Language Pathology*, *11*(2), 163–174. doi:10.1044/1058-0360(2002/016)

Logemann, J. A., & Fisher, H. B. (1981). Vocal tract control in Parkinson's disease. *Journal of Speech and Hearing Disorders*, *46*(4), 348–352. doi:10.1044/jshd.4604.348

Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *The Journal of Speech and Hearing Disorders*, *43*(1), 47–57. doi:10.1044/jshd.4301.47

Logemann, J., Boshes, B., Fisher, H., & Siegfried, J. (1973). The steps in the degeneration of speech and voice control in Parkinson's disease. *Parkinson's Diseases: Rigidity, Akinesia, Behavior*, 101–112.

Lohr, J. B., & Wisniewski, A. A. (1987). *Movement disorders: A neuropsychiatric approach.* Guilford Press.

Ludlow, C. L., & Bassich, C. J. (1984). Relationships between perceptual ratings and acoustic measures of hypokinetic speech. In M. R. McNcil, Rosenbek J. C., & A. E. Aronson (Eds.), *The dysarthrias: Physiology, acoustics, perception, management* (pp. 163–195). San Diego, California: College-Hill Press.

Ludlow, C. L., Connor, N. P., & Bassich, C. J. (1987). Speech timing in Parkinson's and Huntington's disease. *Brain and Language*, *32*(2), 195–214.

Maassen, B. (1986). Marking word boundaries to improve the intelligibility of the speech of the deaf. *Journal of Speech, Language, and Hearing Research*, *29*(2), 227–230.

MacCallum, J. K., Zhang, Y., & Jiang, J. J. (2011). Vowel selection and its effects on perturbation and nonlinear dynamic measures. *Folia Phoniatrica et Logopaedica*, *63*(2), 88–97.

Maniwa, K., Jongman, A., & Wade, T. (2008). Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *123*(2), 1114–1125.

Martel-Sauvageau, V., Macoir, J., Langlois, M., Prud'Homme, M., Cantin, L., & Roy, J.-P. (2014). Changes in vowel articulation with subthalamic nucleus deep brain stimulation in dysarthric speakers with Parkinson's disease. *Parkinson's Disease*, *2014*, 1–9. doi:10.1155/2014/487035

Martel-Sauvageau, V., Roy, J.-P., Cantin, L., Prud'Homme, M., Langlois, M., & Macoir, J. (2015). Articulatory changes in vowel production following STN DBS and levodopa intake in Parkinson's disease. *Parkinson's Disease*, *2015*, 1–7. doi:10.1155/2015/382320

Martens, H., Van Nuffelen, G., Dekens, T., Huici, M. H.-D., Hernández-Díaz, H. A. K., De Letter, M., & De Bodt, M. (2015). The effect of intensive speech rate and intonation therapy on intelligibility in Parkinson's disease. *Journal of Communication Disorders*, *58*, 91–105.

Martínez-Sánchez, F., Meilán, J. J., Carro, J., Íñiguez, C. G., Millian-Morell, L., Valverde, I. M. P., … López, D. E. (2016). Speech rate in Parkinson’s disease: A controlled study. *Neurología (English Edition)*, *31*(7), 466–472. doi:10.1016/j.nrleng.2014.12.014

Maruthy, S., & Raj, N. (2014). Relationship between speech intelligibility and listener effort in Malayalam-speaking individuals with hypokinetic dysarthria. *Speech, Language and Hearing*, *17*(4), 237–245. doi:10.1179/2050571x14z.00000000057

Massen, B., & Povel, D. (1984). The effect of correcting temporal structure on the intelligibility of the deaf. *Speech Communication*, *3*, 123–135.

MATLAB version 9.4.0 (R2018a). (2018). Natick, Massachusetts: The Mathworks, Inc.

McAuliffe, M. J., Fletcher, A. R., Kerr, S. E., O’Beirne, G. A., & Anderson, T. (2017). Effect of dysarthria type, speaking condition, and listener age on speech intelligibility. *American Journal of Speech-Language Pathology*, *26*(1), 113–123. doi:10.1044/2016_ajslp-15-0182

McAuliffe, M. J., Schaefer, M., O’Beirne, G. A., & LaPointe, L. L. (2009). *Effect of noise upon the perception of speech intelligibility in dysarthria*. Retrieved from http://hdl.handle.net/10092/3410

McAuliffe, M. J., Ward, E. C., & Murdoch, B. E. (2006a). Speech production in Parkinson’s disease: I. An electropalatographic investigation of tongue-palate contact patterns. *Clinical Linguistics & Phonetics*, *20*(1), 1–18. doi:10.1080/02699200400001044

McAuliffe, M. J., Ward, E. C., & Murdoch, B. E. (2006b). Speech production in Parkinson’s disease: II. Acoustic and electropalatographic investigation of sentence, word and segment durations. *Clinical Linguistics & Phonetics*, *20*(1), 19–33. doi:10.1080/0269-9200400001069

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of Interspeech 2017* (pp. 498–502). doi:10.21437/interspeech.2017-1386

McCaig, C. M., Adams, S. G., Dykstra, A. D., & Jog, M. (2016). Effect of concurrent walking and interlocutor distance on conversational speech intensity and rate in Parkinson’s disease. *Gait & Posture*, *43*, 132–136.

McCloy, D. (2013). Praat script: Mix speech with noise. GitHub. Retrieved from https://github.com/drammock/praat-semiauto/blob/master/MixSpeechNoise.praat

McHenry, M. (2011). An exploration of listener variability in intelligibility judgments. *American Journal of Speech-Language Pathology*, *20*(2), 119–123. doi:10.1044/1058-0360(2010/10-0059)

McRae, P. A., Tjaden, K., & Schoonings, B. (2002). Acoustic and perceptual consequences of articulatory rate change in Parkinson disease. *Journal of Speech, Language, and Hearing Research*, *45*(1), 35–50. doi:10.1044/1092-4388(2002/003)

Mefferd, A. S., & Corder, E. E. (2014). Assessing articulatory speed performance as a potential factor of slowed speech in older adults. *Journal of Speech, Language, and Hearing Research*, *57*(2), 347–360.

Metter, E. J., & Hanson, W. R. (1986). Clinical and acoustical variability in hypokinetic dysarthria. *Journal of Communication Disorders*, *19*(5), 347–366.

Metz, D. E., Samar, V. J., Schiavetti, N., Sitler, R. W., & Whitehead, R. L. (1985). Acoustic dimensions of hearing-impaired speakers' intelligibility. *Journal of Speech and Hearing Research*, *28*, 345–355.

Midi, I., Dogan, M., Koseoglu, M., Can, G., Sehitoglu, M. A., & Gunal, D. I. (2008). Voice abnormalities and their relation with motor dysfunction in Parkinson's disease. *Acta Neurologica Scandinavica*, *117*(1), 26–34.

Miller, J. L. (1981). Effects of speaking rate on segmental distinctions. In P. D. Elmas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 39–74). Hillsdale, NJ: Lawrence Erlbaum.

Miller, J. L., Green, K. P., & Reeves, A. (1986). Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica*, *43*(1-3), 106–115.

Miller, J. L., O'Rourke, T. B., & Volaitis, L. E. (1997). Internal structure of phonetic categories: Effects of speaking rate. *Phonetica*, *54*(3-4), 121–137. doi:10.1159/000262217

Miller, N. (2013). Measuring up to speech intelligibility. *International Journal of Language & Communication Disorders*, *48*(6), 601–612. doi:10.1111/1460-6984.12061

Miller, N., Allcock, L., Jones, D., Noble, E., Hildreth, A. J., & Burn, D. J. (2007). Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *78*(11), 1188–1190. doi:10.1136/jnnp.2006.110171

Millian-Morell, L., Lopez-Alburquerque, T., Rodriguez-Rodriguez, A., Gomez-Nieto, R., Carro, J., Meilan, J. J., … Lopez, D. E. (2018). Relations between sensorimotor integration and speech disorders in Parkinson's disease. *Current Alzheimer Research*, *15*(2), 149–156.

Monsen, R. B. (1978). Toward measuring how well hearing-impaired children speak. *Journal of Speech and Hearing Research*, *21*, 197–219.

Monsen, R. B. (1983). The oral speech intelligibility of hearing-impaired talkers. *Journal of Speech and Hearing Disorders*, *48*, 286–296. doi:10.1044/jshd.4803.286

Montgomery, E. B. (2007). Deep brain stimulation and speech: A new model of speech function and dysfunction in Parkinson's disease. *Journal of Medical Speech-Language Pathology*, *15*(3), 9–25.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America*, *96*(1), 40–55.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, *47*(5), 1048–1058. doi:10.1044/1092-4388(2004/078)

Mutch, W. J., Strudwick, A., Roy, S. K., & Downie, A. W. (1986). Parkinson's disease: Disability, review, and management. *British Medical Journal*, *293*(6548), 675–677. doi:10.1136/bmj.293.6548.675

Müller, J., Wenning, G. K., Verny, M., McKee, A., Chaudhuri, K. R., Jellinger, K., … Litvan, I. (2001). Progression of dysarthria and dysphagia in postmortem-confirmed parkinsonian disorders. *Archives of Neurology*, *58*(2), 259–264. doi:10.1001/archneur.58.2.259

Neel, A. (2009). Effects of loud and amplified speech on sentence and word intelligibility in Parkinson disease. *Journal of Speech, Language, and Hearing Research*.

Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research*, *51*(3), 574–585.

Netsell, R., Daniel, B., & Celesia, G. G. (1975). Acceleration and weakness in parkinsonian dysarthria. *Journal of Speech and Hearing Disorders*, *40*(2), 170–178.

Nickerson, R. S., & Stevens, K. N. (1980). Approaches to the study of the relationship between intelligibility and physical properties of speech. In J. Subtelny (Ed.), *Speech assessment and speech improvement for the hearing impaired* (pp. 338–364). A. G. Bell Association for the Deaf.

Nip, I. S., & Green, J. R. (2013). Increases in cognitive and linguistic processing primarily account for increases in speaking rate with age. *Child Development*, *84*(4), 1324–1337.

Nishio, M., Tanaka, Y., Sakabibara, C., & Abe, N. (2011). Effectiveness of speech rate-conversion software for patients with dysarthria. *Journal of Communications Research*, *2*(1).

Nussbaum, R. L., & Ellis, C. E. (2003). Alzheimer's disease and Parkinson's disease. *New England Journal of Medicine*, *348*(14), 1356–1364. doi:10.1056/nejm2003ra020003

Obeso, J. A., Rodriguez-Oroz, M. C., Goetz, C. G., Marin, C., Kordower, J. H., Rodriguez, M., … Halliday, G. (2010). Missing pieces in the Parkinson's disease puzzle. *Nature Medicine*, *16*(6), 653. doi:10.1038/nm.2165

Odlozinski, L. M. (1998). *An acoustic analysis of speech rate control procedures in Parkinson's disease* (Master's thesis). University of Western Ontario.

Oguz, H., Tunc, T., Safak, M. A., Inan, L., Kargin, S., & Demirci, M. (2006). Objective voice changes in nondysphonic Parkinson's disease patients. *Journal of Otolaryngology*, *35*(5).

Okun, M. S. (2012). Deep-brain stimulation for Parkinson's disease. *New England Journal of Medicine*, *367*(16), 1529–1538.

Okun, M. S., & Foote, K. D. (2004). A mnemonic for Parkinson disease patients considering DBS: A tool to improve perceived outcome of surgery. *The Neurologist*, *10*(5), 290–290. doi:10.1097/01.nrl.0000138737.97544.7c

Osberger, M. J., & Levitt, H. (1979). The effect of timing errors on the intelligibility of deaf children's speech. *The Journal of the Acoustical Society of America*, *66*(5), 1316–1324.

Özsancak, C., Auzou, P., Jan, M., & Hannequin, D. (2001). Measurement of voice onset time in dysarthric patients: Methodological considerations. *Folia Phoniatrica et Logopaedica*, *53*(1), 48–57.

Parveen, S., & Goberman, A. M. (2014). Presence of stop bursts and multiple bursts in individuals with Parkinson disease. *International Journal of Speech-Language Pathology*, *16*(5), 456–463. doi:10.3109/17549507.2013.808702

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research*, *28*(1), 96–103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing ii: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, *29*(4), 434–446.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing iii: An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research*, *32*(3), 600–603.

Pilon, M. A., McIntosh, K. W., & Thaut, M. H. (1998). Auditory vs visual speech timing cues as external rate control to enhance verbal intelligibility in mixed spastic ataxic dysarthric speakers: A pilot study. *Brain Injury*, *12*(9), 793–803.

Plaha, P., Ben-Shlomo, Y., Patel, N. K., & Gill, S. S. (2006). Stimulation of the caudal zona incerta is superior to stimulation of the subthalamic nucleus in improving contralateral parkinsonism. *Brain*, *129*(7), 1732–1747.

Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., … Lang, A. E. (2017). Parkinson disease. *Nature Reviews Disease Primers*, *3*(1), 17013. doi:10.1038/nrdp.2017.14

Postuma, R. B. (2014). Prodromal Parkinson's disease–using REM sleep behavior disorder as a window. *Parkinsonism & Related Disorders*, *20*, S1–S4.

Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., … Deuschl, G. (2015). MDS clinical diagnostic criteria for Parkinson's disease. *Movement Disorders*, *30*(12), 1591–1601. doi:10.1002/mds.26424

Putzer, M., Barry, W. J., & Moringlane, J. R. (2008). Effect of bilateral stimulation of the subthalamic nucleus on different speech subsystems in patients with Parkinson's disease. *Clinical Linguistics and Phonetics*, *22*(12), 957–973.

Quené, H. (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *The Journal of the Acoustical Society of America*, *123*(2), 1104–1113.

Ramig, L. A. (1983). Effects of physiological aging on speaking and reading rates. *Journal of Communication Disorders*, *16*(3), 217–226.

Ramig, L. A., & Ringel, R. L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. *Journal of Speech, Language, and Hearing Research*, *26*(1), 22–30.

Ramig, L. A., Titze, I. R., Scherer, R. C., & Ringel, S. P. (1988). Acoustic analysis of voices of patients with neurologic disease: Rationale and preliminary data. *Annals of Otology, Rhinology & Laryngology*, *97*(2), 164–172.

Ramig, L. O., & Dromey, C. (1996). Aerodynamic mechanisms underlying treatment-related changes in vocal intensity in patients with Parkinson disease. *Journal of Speech, Language, and Hearing Research*, *39*(4), 798–807. doi:10.1044/jshr.3904.798

Ramig, L. O., Fox, C., & Sapir, S. (2004). Parkinson's disease: Speech and voice disorders and their treatment with the Lee Silverman Voice Treatment. In *Seminars in speech and language* (Vol. 25, pp. 169–180). Copyright© 2004 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA.

Rascol, O., Payoux, P., Ory, F., Ferreira, J. J., Brefel-Courbon, C., & Montastruc, J.-L. (2003). Limitations of current Parkinson's disease therapy. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *53*(S3), S3–S15.

Ravizza, S. M. (2003). Dissociating the performance of cortical and subcortical patients on phonemic tasks. *Brain and Cognition*, *53*(2), 301–310. doi:10.1016/s0278-2626(03)00131-3

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Read, J., Miller, N., & Kitsou, N. (2018). Is there an order of loss of sounds in speakers with Parkinson's disease? *Clinical Linguistics & Phonetics*, *32*(11), 997–1011.

Reijnders, J. S., Ehrt, U., Weber, W. E., Aarsland, D., & Leentjens, A. F. (2008). A systematic review of prevalence studies of depression in Parkinson's disease. *Movement Disorders*, *23*(2), 183–189.

Robertson, L. T., St George, R. J., Carlson-Kuhta, P., Hogarth, P., Burchiel, K. J., & Horak, F. B. (2011). Site of deep brain stimulation and jaw velocity in Parkinson disease. *Journal of Neurosurgery*, *115*(5), 985–994.

Rogers, H. (2014). *The sounds of language: An introduction to phonetics*. Routledge.

Rosen, K. M., Kent, R. D., Delaney, A. L., & Duffy, J. R. (2006). Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers.

Journal of Speech, Language, and Hearing Research, *49*(2), 395–411. doi:10.1044/1092-4388(2006/031)

Rousseaux, M., Krystkowiak, P., Kozlowski, O., Özsancak, C., Blond, S., & Destée, A. (2004). Effects of subthalamic nucleus stimulation on parkinsonian dysarthria and speech intelligibility. *Journal of Neurology*, *251*(3), 327–334. doi:10.1007/s00415-004-0327-1

Roy, N., Nissen, S. L., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal therapy. *Journal of Communication Disorders*, *42*(2), 124–135.

Rusz, J., Čmejla, R., Ruzickova, H., & Ruzicka, E. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *The Journal of the Acoustical Society of America*, *129*(1), 350–367.

Rusz, J., Čmejla, R., Tykalova, T., Ruzickova, H., Klempir, J., Majerova, V., … Ruzicka, E. (2013). Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *The Journal of the Acoustical Society of America*, *134*(3), 2171–2181. doi:10.1121/1.4816541

Rusz, J., Čmejla, R., Růžičková, H., Klempíř, J., Majerová, V., Picmausová, J., … Růžička, E. (2013). Evaluation of speech impairment in early stages of Parkinson's disease: A prospective study with the role of pharmacotherapy. *Journal of Neural Transmission*, *120*(2), 319–329.

Ryalls, J. H., & Lieberman, P. (1982). Fundamental frequency and vowel perception. *The Journal of the Acoustical Society of America*, *72*(5), 1631–1634. doi:10.1121/1.388499

Santens, P., De Letter, M., Van Borsel, J., De Reuck, J., & Caemaert, J. (2003). Lateralized effects of subthalamic nucleus stimulation on different aspects of speech in Parkinson's disease. *Brain and Language*, *87*(2), 253–258.

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*, *57*(4), 1330–1343.

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech. *Journal of Speech, Language, and Hearing Research*, *53*(1), 114–125. doi:10.1044/1092-4388(2009/08-0184)

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2011). Acoustic metrics of vowel articulation in Parkinson's disease: Vowel space area (VSA) vs. Vowel articulation index (VAI). In C. Manfredi (Ed.), *Models and analysis of vocal emissions for biomedical applications* (pp. 173–175).

Sapir, S., Spielman, J., Ramig, L., Story, B., & Fox, C. (2007). Effects of intensive voice treatment LSVT on vowel articulation in dysarthric individuals with idiopathic parkinson disease: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research*, *50*(4), 899–912. doi:10.1044/1092-4388(2007/064)

Schulz, G., Hosey, L., Bradberry, T., Stager, S., Lee, L., Pawha, R., … Braun, A. (2012). Selective left, right and bilateral stimulation of subthalamic nuclei in Parkinson's disease: Differential effects on motor, speech and language function. *Journal of Parkinson's Disease*, *2*(1), 29–40. doi:10.3233/JPD-2012-11049

Sidtis, D., Cameron, K., Bonura, L., & Sidtis, J. (2012). Speech intelligibility by listening in Parkinson speech with and without deep brain stimulation: Task effects. *Journal of Neurolinguistics*, *25*(2), 121–132. doi:10.1016/j.jneuroling.2011.08.004

Sidtis, D. V. L., Rogers, T., Godier, V., Tagliati, M., & Sidtis, J. J. (2010). Voice and fluency changes as a function of speech task and deep brain stimulation. *Journal of Speech, Language, and Hearing Research*, *53*(5), 1167–1177. doi:10.1044/1092-4388(2010/09-0154)

Sidtis, J. J., Alken, A. G., Tagliati, M., Alterman, R., & Van Lancker Sidtis, D. (2016). Subthalamic stimulation reduces vowel space at the initiation of sustained production: Implications for articulatory motor control in Parkinson's disease. *Journal of Alzheimer's Disease*, *6*(2), 361–370. doi:10.3233/jpd-150739

Silva, L. F. e, Gama, A. C. C., Cardoso, F. E. C., Reis, C. A. da C., & Bassi, I. B. (2012). Idiopathic Parkinson's disease: Vocal and quality of life analysis. *Arquivos de Neuro-Psiquiatria*, *70*(9), 674–679.

Simpson, A., & Ericsdotter, C. (2007). Sex-specific differences in F0 and vowel space. In *XVIth international congress of phonetic sciences* (pp. 933–936).

Skodda, S. (2011). Aspects of speech rate and regularity in Parkinson's disease. *Journal of the Neurological Sciences*, *310*(1-2), 231–236. doi:10.1016/j.jns.2011.07.020

Skodda, S., Grönheit, W., & Schlegel, U. (2012). Impairment of vowel articulation as a possible marker of disease progression in Parkinson's disease. *PloS One*, *7*(2), e32132. doi:10.1371/journal.pone.0032132

Skodda, S., Grönheit, W., Schlegel, U., Südmeyer, M., Schnitzler, A., & Wojtecki, L. (2013). Effect of subthalamic stimulation on voice and speech in Parkinson's disease: For the better or worse? *Frontiers in Neurology*, *4*.

Skodda, S., & Schlegel, U. (2008). Speech rate and rhythm in Parkinson's disease. *Movement Disorders*, *23*, 985–992. doi:10.1002/mds.21996

Skodda, S., Visser, W., & Schlegel, U. (2011). Vowel articulation in Parkinson's disease. *Journal of Voice*, *25*, 467–472. doi:10.1016/j.jvoice.2010.01.009

Smith, B. L., Wasowicz, J., & Preston, J. (1987). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech, Language, and Hearing Research*, *30*(4), 522–529.

Smith, B. L., Wasowicz, J., & Preston, J. (1988). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech and Hearing Research*, *30*, 522–529.

Spencer, K. A., Morgan, K. W., & Blond, E. (2009). Dopaminergic medication effects on the speech of individuals with Parkinson's disease. *Journal of Medical Speech-Language Pathology*, *17*(3), 125–145.

Sperry, E. E., & Klich, R. J. (1992). Speech breathing in senescent and younger women during oral reading. *Journal of Speech, Language, and Hearing Research*, *35*(6), 1246–1255.

Spielman, J., Mahler, L., Halpern, A., Gilley, P., Klepitskaya, O., & Ramig, L. (2011). Intensive voice treatment (LSVTLOUD) for Parkinson's disease following deep brain stimulation of the subthalamic nucleus. *Journal of Communication Disorders*, *44*(6), 688–700. doi:10.1016/j.jcomdis.2011.05.003

Spitzer, S., Liss, J., Caviness, J. N., & Adler, C. (2000). An exploration of familiarization effects in the perception of hypokinetic and ataxic dysarthric speech. *Journal of Medical Speech-Language Pathology*, *8*(4), 285–293.

Stevens, K., Nickerson, R., & Rollins, A. (1983). Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training, and personnel preparation* (pp. 35–51). Baltimore, MD: University Park Press.

Stipancic, K. L., Tjaden, K., & Wilding, G. (2016). Comparison of intelligibility measures for adults with Parkinson's disease, adults with multiple sclerosis, and healthy controls. *Journal of Speech, Language, and Hearing Research*, *59*(2), 230–238. doi:10.1044/2015_jslhr-s-15-0271

Stuart-Smith, J., Sonderegger, M., Rathcke, T., & Macdonald, R. (2015). The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. *Laboratory Phonology*, *6*(3-4), 505–549.

Summerfield, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(5), 1074.

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *The Journal of the Acoustical Society of America*, *84*(3), 917–928. doi:10.1121/1.396660

Sussman, J. E., & Tjaden, K. (2012). Perceptual measures of speech from individuals with Parkinson's disease and multiple sclerosis: Intelligibility and beyond. *Journal of Speech, Language, and Hearing Research*, *55*(4), 1208–1219. doi:10.1044/1092-4388(2011/11-0048)

Tanaka, Y., Nishio, M., & Niimi, S. (2011). Vocal acoustic characteristics of patients with Parkinson's disease. *Folia Phoniatrica et Logopaedica*, *63*(5), 223–230.

Tanaka, Y., Tsuboi, T., Watanabe, H., Kajita, Y., Fujimoto, Y., Ohdake, R., … others. (2015). Voice features of Parkinson's disease patients with subthalamic nucleus deep brain stimulation. *Journal of Neurology*, *262*(5), 1173–1181. doi:10.1007/s00415-015-7681-z

Tanaka, Y., Tsuboi, T., Watanabe, H., Kajita, Y., Nakatsubo, D., Fujimoto, Y., … others. (2016). Articulation features of Parkinson's disease patients with subthalamic nucleus deep brain stimulation. *Journal of Parkinson's Disease*, *6*(4), 811–819. doi:10.3233/jpd-160838

Tanner, J., Sonderegger, M., & Wagner, M. (2017). Production planning and coronal stop deletion in spontaneous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *8*(1).

Tetrud, J. W. (1991). Preclinical Parkinson's disease: Detection of motor and nonmotor manifestations. *Neurology*, *41*(Issue 5, Supplement 2), 69–71. doi:10.1212/wnl.41.5_suppl_2.69

Thobois, S., Prange, S., Scheiber, C., & Broussolle, E. (2019). What a neurologist should know about PET and SPECT functional imaging for parkinsonism: A practical perspective. *Parkinsonism & Related Disorders*, *59*, 93–100. doi:10.1016/j.parkreldis.2018.08.016

Tjaden, K. (2000a). An acoustic study of coarticulation in dysarthric speakers with parkinson disease. *Journal of Speech, Language, and Hearing Research*, *43*, 1466–1480.

Tjaden, K. (2000b). A preliminary study of factors influencing perception of articulatory rate in Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *43*, 997–101.

Tjaden, K. (2003). Anticipatory coarticulation in multiple sclerosis and Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *46*(4), 990–1008. doi:10.1044/1092-4388(2003/077)

Tjaden, K., Greenlee, J., Brenk, F. J. van, Silverman, S., & Corcos, D. (2018). Perceptual and acoustic correlates of DBS of subthalamic nucleus versus globus pallidus interna for IPD: A comparative pilot study. In *19th biennial conference on motor speech*.

Tjaden, K., Lam, J., & Wilding, G. (2013a). Vowel acoustics in Parkinson's disease and multiple sclerosis: Comparison of clear, loud, and slow speaking conditions. *Journal of*

Speech, Language, and Hearing Research, *56*(5), 1485–1502. doi:10.1044/1092-4388(2013/12-0259)

Tjaden, K., & Martel-Sauvageau, V. (2017). Consonant acoustics in Parkinson's disease and multiple sclerosis: Comparison of clear and loud speaking conditions. *American Journal of Speech-Language Pathology*, *26*(2S), 569–582.

Tjaden, K., Richards, E., Kuo, C., Wilding, G., & Sussman, J. (2013b). Acoustic and perceptual consequences of clear and loud speech. *Folia Phoniatrica et Logopaedica*, *65*(4), 214–220. doi:10.1159/000355867

Tjaden, K., Richards, E., Kuo, C., Wilding, G., & Sussman, J. (2014a). Acoustic and perceptual consequences of clear and loud speech. *Folia Phoniatrica et Logopaedica*, *65*(4), 214–220.

Tjaden, K., Rivera, D., Wilding, G., & Turner, G. S. (2005). Characteristics of the lax vowel space in dysarthria. *Journal of Speech, Language, and Hearing Research*, *48*(3), 554–566. doi:10.1044/1092-4388(2005/038)

Tjaden, K., & Sussman, J. (2006). Perception of coarticulatory information in normal speech and dysarthria. *Journal of Speech, Language, and Hearing Research*, *49*(4), 888–902. doi:10.1044/1092-4388(2006/064)

Tjaden, K., Sussman, J. E., & Wilding, G. E. (2014b). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research*, *57*, 779–792. doi:10.1044/2014_JSLHR-S-12-0372

Tjaden, K., & Weismer, G. (1998). Speaking-rate-induced variability in F2 trajectories. *Journal of Speech, Language, and Hearing Research*, *41*(5), 976–989.

Tjaden, K., & Wilding, G. (2011a). Effects of speaking task on intelligibility in Parkinson's disease. *Clinical Linguistics & Phonetics*, *25*(2), 155–168. doi:10.3109/02699206.2010.520185

Tjaden, K., & Wilding, G. (2011b). Speech and pause characteristics associated with voluntary rate reduction in Parkinson's disease and Multiple Sclerosis. *Journal of Communication Disorders*, *44*, 655–665. doi:10.1016/j.jcomdis.2011.06.003

Tjaden, K., & Wilding, G. (2011c). The impact of rate reduction and increased loudness on fundamental frequency characteristics in dysarthria. *Folia Phoniatrica et Logopaedica*, *63*(4), 178–186. doi:10.1159/000316315

Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research*, *47*(4), 766–783. doi:10.1044/1092-4388(2004/058)

Törnqvist, A. L., Schalén, L., & Rehncrona, S. (2005). Effects of different electrical parameter settings on the intelligibility of speech in patients with Parkinson's disease treated with subthalamic deep brain stimulation. *Movement Disorders*, *20*(4), 416–423. doi:10.1002/mds.20348

Tripoliti, E., Limousin, P., Foltynie, T., Candelario, J., Aviles-Olmos, I., Hariz, M. I., & Zrinzo, L. (2014). Predictive factors of speech intelligibility following subthalamic nucleus stimulation in consecutive patients with Parkinson's disease. *Movement Disorders*, *29*(4), 532–538. doi:10.1002/mds.25816

Tripoliti, E., Zrinzo, L., Martinez-Torres, I., Tisch, S., Frost, E., Borrell, E., … Limousin, P. (2008). Effects of contact location and voltage amplitude on speech and movement in bilateral subthalamic nucleus deep brain stimulation. *Movement Disorders*, *23*(16), 2377–2383. doi:10.1002/mds.22296

Tripoliti, E., Zrinzo, L., & Martinez-Torresetal, I. (2011). Effects of sub-thalamic stimulation on speech of consecutive patients with Parkinson disease. *Neurology*, *76*(1), 80–86. doi:10.1212/wnl.0b013e318203e7d0

Tsao, Y.-C., & Iqbal, K. (2006). Can acoustic vowel space predict the habitual speech rate of the speaker? In *27th annual international conference of the engineering in medicine and biology society, 2005. IEEE-EMBS 2005.* (pp. 1220–1223). IEEE.

Tsao, Y.-C., & Weismer, G. (1997). Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech, Language, and Hearing Research*, *40*(4), 858–866.

Tsao, Y.-C., Weismer, G., & Iqbal, K. (2006). Interspeaker variation in habitual speaking rate: Additional evidence. *Journal of Speech, Language, and Hearing Research*, *49*(5), 1156–1164. doi:10.1044/1092-4388(2006/083)

Tsuboi, T., Watanabe, H., Tanaka, Y., Ohdake, R., Yoneyama, N., Hara, K., … others. (2014). Distinct phenotypes of speech and voice disorders in Parkinson's disease after subthalamic nucleus deep brain stimulation. *Journal of Neurology, Neurosurgery & Psychiatry*, *86*(8), 856–864. doi:10.1136/jnnp-2014-308043

Turner, G., Martin, H., & de Jonge, R. (2008). Influence of amplification on word intelligibility and phonetic contrast errors in dysarthria. *Journal of Medical Speech Language Pathology*.

Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, *38*(5), 1001–1013. doi:10.1044/jshr.3805.1001

Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech, Language, and Hearing Research*, *39*(3), 494–509. doi:10.1044/jshr.3903.494

van Brenk, F., Terband, H., van Lieshout, P., Lowit, A., & Maassen, B. (2009). An analysis of speech rate strategies in ageing. *Interspeech*, 792–795.

Van Nuffelen, G., De Bodt, M., Vanderwegen, J., Van de Heyning, P., & Wuyts, F. (2010). Effect of rate control on speech production and intelligibility in dysarthria. *Folia Phoniatrica et Logopaedica*, *62*(3), 110–119.

Van Nuffelen, G., De Bodt, M., Wuyts, F., & Van de Heyning, P. (2009). The effect of rate control on speech rate and intelligibility of dysarthric speech. *Folia Phoniatrica et Logopaedica*, *61*(2), 69–75. doi:10.1159/000208805

Van Son, R. J., & Pols, L. C. (1990). Formant frequencies of Dutch vowels in a text, read at normal and fast rate. *The Journal of the Acoustical Society of America*, *88*(4), 1683–1693. doi:10.1121/1.400243

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech*, *47*(3), 297–308.

Vingerhoets, F. J., Villemure, J.-G., Temperli, P., Pollo, C., Pralong, E., & Ghika, J. (2002). Subthalamic DBS replaces levodopa in Parkinson's disease two-year follow-up. *Neurology*, *58*(3), 396–401. doi:10.1212/wnl.58.3.396

Vogel, A. P., Rommel, N., Oettinger, A., Horger, M., Krumm, P., Kraus, E.-M., … Synofzik, M. (2017). Speech and swallowing abnormalities in adults with POLG associated ataxia (POLG-A). *Mitochondrion*.

Wagner, M. (2018). ProsodyLab Experimenter. GitHub. Retrieved from https://github.com/prosodylab/prosodylab-experimenter

Walsh, B., & Smith, A. (2011). Linguistic complexity, speech production, and comprehension in Parkinson's disease: Behavioral and physiological indices. *Journal of Speech, Language, and Hearing Research*, *54*, 787–802.

Walsh, B., & Smith, A. (2012). Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease. *Movement Disorders*, *27*(7), 843–850. doi:10.1002/mds.24888

Wang, E. Q., Metman, L. V., Bakay, R. A., Arzbaecher, J., Bernard, B., & Corcos, D. M. (2006). Hemisphere-specific effects of subthalamic nucleus deep brain stimulation on speaking rate and articulatory accuracy of syllable repetitions in Parkinson's disease. *Journal of Medical Speech-Language Pathology*, *14*(4), 323.

Wang, Y.-t., Kent, R. D., Duffy, J. R., Thomas, J. E., & Weismer, G. (2004). Alternating motion rate as an index of speech motor disorder in traumatic brain injury. *Clinical Linguistics & Phonetics*, *18*(1), 57–84. doi:10.1080/02699200310001596160

Watson, P. J., & Munson, B. (2008). Parkinson's disease and the effect of lexical factors on vowel articulation. *The Journal of the Acoustical Society of America*, *124*(5), EL291–EL295. doi:10.1121/1.2987464

Weir-Mayta, P., Spencer, K. A., Eadie, T. L., Yorkston, K., Savaglio, S., & Woollcott, C. (2017). Internally versus externally cued speech in Parkinson's disease and cerebellar disease. *American Journal of Speech-Language Pathology*, *26*(2S), 583–595.

Weismer, G. (1984a). Acoustic descriptions of dysarthric speech: Perceptual correlates and physiological inferences. In *Seminars in speech and language* (Vol. 5, pp. 293–314). © 1984 by Thieme Medical Publishers, Inc.

Weismer, G. (1984b). Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination. In M. McNeil (Ed.), *The dysarthrias: Physiology, acoustics, perception, management* (pp. 101–130). College Hill Press San Diego.

Weismer, G. (1991). Assessment of articulatory timing. *NIDCD Monograph*, *1*, 84–95.

Weismer, G. (2006). Philosophy of research in motor speech disorders. *Clinical Linguistics & Phonetics*, *20*(5), 315–349. doi:10.1080/02699200400024806

Weismer, G. (2008). Speech intelligibility. In M. J. Ball, M. R. Perkins, N. Muller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 568–582). Oxford, England: Blackwell.

Weismer, G., Jeng, J.-Y., Laures, J. S., Kent, R. D., & Kent, J. F. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica*, *53*(1), 1–18. doi:10.1159/000052649

Weismer, G., & Laures, J. S. (2002). Direct magnitude estimates of speech intelligibility in dysarthria: Effects of a chosen standard. *Journal of Speech, Language, and Hearing Research*, *45*(3), 421–433. doi:10.1044/1092-4388(2002/033)

Weismer, G., Laures, J. S., Jeng, J.-Y., Kent, R. D., & Kent, J. F. (2000). Effect of speaking rate manipulations on acoustic and perceptual aspects of the dysarthria in amyotrophic lateral sclerosis. *Folia Phoniatrica et Logopaedica*, *52*(5), 201–219. doi:10.1159/000021536

Weismer, G., & Martin, R. E. (1992). Acoustic and perceptual approaches to the study of intelligibility. In R. D. Kent (Ed.), *Intelligibility in speech disorders: Theory, measurement and management* (pp. 67–118). John Benjamins Publishing Company. doi:10.1075/sspcl.1.04wei

Weismer, G., Martin, R., Kent, R. D., & Kent, J. F. (1992). Formant trajectory characteristics of males with amyotrophic lateral sclerosis. *The Journal of the Acoustical Society of America*, *91*(2), 1085–1098.

Weismer, G., Yunusova, Y., & Bunton, K. (2012). Measures to evaluate the effects of DBS on speech production. *Journal of Neurolinguistics*, *25*(2), 74–94. doi:10.1016/j.jneuroling.2011.08.006

Whitehill, T. L., Ciocca, V., & Yiu, E. M.-L. (2004). Perceptual and acoustic predictors of intelligibility and acceptability in Cantonese speakers with dysarthria. *Journal of Medical Speech-Language Pathology*, *12*(4), 229–234.

Whitfield, J. A., & Goberman, A. M. (2014). Articulatory-acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. *Journal of Communication Disorders*, *51*, 19–28.

Winkworth, A. L., Davis, P. J., Ellis, E., & Adams, R. D. (1994). Variability and consistency in speech breathing during reading: Lung volumes, speech intensity, and linguistic factors. *Journal of Speech, Language, and Hearing Research*, *37*(3), 535–556.

Wirdefeldt, K., Adami, H.-O., Cole, P., Trichopoulos, D., & Mandel, J. (2011). Epidemiology and etiology of Parkinson's disease: A review of the evidence. *European Journal of Epidemiology*, *26*(1), 1–58. doi:10.1007/s10654-011-9581-6

Wohlert, A. B., & Hammen, V. L. (2000). Lip muscle activity related to speech rate and loudness. *Journal of Speech, Language, and Hearing Research*, *43*(5), 1229–1239. doi:10.1044/jslhr.4305.1229

Wohlert, A. B., & Smith, A. (1998). Spatiotemporal stability of lip movements in older adult speakers. *Journal of Speech, Language, and Hearing Research*, *41*(1), 41–50.

Wong, M. N., Murdoch, B. E., & Whelan, B.-M. (2011). Lingual kinematics in dysarthric and nondysarthric speakers with Parkinson's disease. *Parkinson's Disease*, *2011*, 1–8. doi:10.4061/2011/352838

Xie, X., & Fowler, C. A. (2013). Listening with a foreign-accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English. *Journal of Phonetics*, *41*(5), 369–378.

Xie, Y., Zhang, Y., Zheng, Z., Liu, A., Wang, X., Zhuang, P., … Wang, X. (2011). Changes in speech characters of patients with Parkinson's disease after bilateral subthalamic nucleus stimulation. *Journal of Voice*, *25*(6), 751–758.

Yoho, S. E., & Borrie, S. A. (2018). Combining degradations: The effect of background noise on intelligibility of disordered speech. *The Journal of the Acoustical Society of America*, *143*(1), 281–286. doi:10.1121/1.5021254

Yorkston, K. M., & Beukelman, D. R. (1978). A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of Communication Disorders*, *11*(6), 499–512. doi:10.1016/0021-9924(78)90024-2

Yorkston, K. M., & Beukelman, D. R. (1981a). *Assessment of intelligibility of dysarthric speech*. (P. Version, Ed.). Pro-Ed.

Yorkston, K. M., & Beukelman, D. R. (1981b). Ataxic dysarthria: treatment sequences based on intelligibility and prosodic considerations. *Journal of Speech and Hearing Disorders*, *46*(4), 398–404.

Yorkston, K. M., & Beukelman, D. R. (1983). The influence of judge familiarization with the speaker on dysarthric speech intelligibility. *Clinical Dysarthria*, 155–163.

Yorkston, K. M., Beukelman, D. R., & Bell, K. R. (1988). *Clinical management of dysarthric speakers*. (J. C. Rosenbek, Ed.). College Hill.

Yorkston, K. M., Beukelman, D. R., Strand, E. A., & Bell, K. R. (1999). *Management of motor speech disorders in children and adults* (2nd ed.). Austin: Pro-Ed.

Yorkston, K. M., Beukelman, D. R., & Tice, R. (1996a). Sentence Intelligibility Test. *Communication Disorders Software. Distributed by the Institute for Rehabilitation Science and Engineering at Madonna Rehabilitation Hospital, Lincoln, NE.*

Yorkston, K. M., Dowden, P. A., & Beukelman, D. R. (1992). Intelligibility measurement as a tool in the clinical management of dysarthric speakers. In *Intelligibility in speech disorders: Theory, measurement and management* (pp. 265–286). John Benjamins Publishing Company.

Yorkston, K. M., Hakel, M., Beukelman, D. R., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology*, *15*(2).

Yorkston, K. M., Hammen, V. L., Beukelman, D. R., & Traynor, C. D. (1990). The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders*, *55*(3), 550–560. doi:10.1044/jshd.5503.550

Yorkston, K. M., Strand, E. A., & Kennedy, M. R. T. (1996b). Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology*, *5*, 55–66. doi:10.1044/1058-0360.0501.55

Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation. In *INTERSPEECH*. 196 wpm: average conversational speaking rate in healthy adults.

Yunusova, Y., Weismer, G., Westbury, J. R., & Lindstrom, M. J. (2008). Articulatory movements during vowels in speakers with dysarthria and healthy controls. *Journal of Speech, Language, and Hearing Research*, *51*(3), 596–611. doi:10.1044/1092-4388(2008/043)

Yunusova, Y., Westbury, J., & Weismer, G. (2005). Articulatory movements during vowels produced by speakers with dysarthria and normal controls. *The Journal of the Acoustical Society of America*, *117*(4), 2605–2605.

Yücetürk, A., Yılmaz, H., Eğrilmez, M., & Karaca, S. (2002). Voice analysis and videolaryngostroboscopy in patients with Parkinson's disease. *European Archives of Oto-Rhino-Laryngology*, *259*(6), 290–293.

Zhou, X. P., Lee, V. S., Wang, E. Q., & Jiang, J. J. (2009). Evaluation of the effects of deep brain stimulation of the subthalamic nucleus and levodopa treatment on parkinsonian voice using perturbation, nonlinear dynamic, and perceptual analysis. *Folia Phoniatrica et Logopaedica*, *61*(4), 189–199. doi:10.1159/000227996

Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, *22*(2), 121–140.

Vita

Name:	Thea Knowles
Post-secondary Education and Degrees:	McGill University Montreal, Quebec, Canada 2008 – 2012 BA Honours Western University London, Ontario, Canada 2014 – 2019 MClSc/PhD
Selected Honours and Awards:	Ontario Graduate Scholarship 2017-2018, 2018-2019 Parkinson Society Southwestern Ontario Graduate Student Research Program Award 2017-2018 Western University "3 Minute Thesis" First Place Winner 2016 Harmonize for Speech Graduate Student Award 2016 Ontario Association of Speech-Language Pathologists and Audiologists Student Achievement Award 2015
Related Work Experience:	Teaching Assistant, Western University 2016-2018 Research Assistant, Western University 2014-2019
Selected Publications:	Knowles, T., Clayards, M., & Sonderegger, M. (2018). Examining factors influencing the viability of automatic acoustic analysis of child speech. Journal of Speech, Language, and Hearing Research, 61(10), 2487-2501. doi: 10.1044/2018 JSLHR-S-17-0275. Knowles, T., Adams, S., Abeyesekera, A., Mancinelli, C., Gilmore, G., & Jog, M. (2018). Deep brain stimulation of the subthalamic nucleus parameter optimization for vowel acoustics and speech intelligibility in Parkinson's disease. Journal of Speech, Language, and Hearing Research, 61(3), 510- 524. doi: 10.1044/2017 JSLHR-S-17-0157. Knowles, T., Clayards, M., Sonderegger, M., Wagner, M., Nadig, A., & Onishi, K. (2015). Automatic forced alignment on child speech: Directions for improvement. Proceedings of Meetings on Acoustics. doi: 25:060001.