NEURAL CORRELATES OF PHONETIC-INTENSITY ENCODING

by

Christina Angela Dastolfo-Hromack

Bachelor of Arts, Allegheny College, 2008

Master of Science, University of Pittsburgh, 2011

Submitted to the Graduate Faculty of the

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Christina Angela Dastolfo-Hromack

It was defended on

December 16, 2021

and approved by

Robert Turner, PhD
Professor, Department of Neurobiology
Dietrich School of Arts & Sciences, University of Pittsburgh

Jason Bohland, PhD
Assistant Professor, Department of Communication Science and Disorders
School of Health and Rehabilitation Sciences, University of Pittsburgh

Bharath Chandrasekaran, PhD
Professor, Department of Communication Science and Disorders
School of Health and Rehabilitation Sciences, University of Pittsburgh

Dissertation Director: Susan Shaiman, PhD, CCC-SLP
Associate Professor, Department of Communication Science and Disorders
School of Health and Rehabilitation Sciences, University of Pittsburgh

Neural Correlates of Phonetic-Intensity Encoding

Christina Angela Dastolfo-Hromack, PhD

University of Pittsburgh, 2021

Speech therapy for patients with Parkinson's disease (PWPD) capitalizes on a physiologic link between vocal intensity and articulation (termed here as 'phonetic-intensity encoding'), causing a patient's articulation to improve given the single objective to speak loudly. However, because specific neurological correlates of this effect remain unknown, phonetic-intensity encoding is not described in models of speech motor control. This project investigated the neurological origins of phonetic-intensity encoding within the cortico-basal ganglia loop (precentral gyrus and subthalamic nucleus) using electrophysiologic recordings gathered from patients undergoing deep brain stimulation implantation surgery. Patients were asked to speak nonsense syllables loudly and softly during the awake portion of surgery. Electrophysiologic signals were decomposed into theta band power (4-8 Hz) and gamma band power (70-150 Hz) which were used as the dependent variable in a mixed-effects model, predicted by phoneme identity, vocal intensity and electrode location. Results showed that phonetic-intensity encoding occurs as vocal intensity-dependent increases in theta band power at phoneme-specific locations with the subthalamic nucleus. Simultaneous phonemic and vocal intensity-dependent changes in neural power were observed in precentral gyrus gamma band power, although no interaction between the variables was found. Secondary analysis concluded that both the precentral gyrus and subthalamic nucleus power predict changes in vocal intensity production, although averaged gamma band power displayed a negative association with intensity, while theta band power displayed a positive association. These results confirm hypotheses that the basal ganglia-cortical

loop is involved in vocal intensity processing and the subthalamic nucleus is an important node in phonetic-intensity encoding.

Table of Contents

List of Tables

List of Figures

Preface

This dissertation would not have been possible without extraordinary support and guidance from many sources. The journey through a PhD program is challenging, and I am thankful for the generosity and kindness which allowed me to persevere through challenges.

First, I need to thank my husband, Bobby Hromack, who has been a continuous source of support through this entire process. He has endured the challenges and triumphs at every turn and has always instilled confidence in my abilities, even during challenging times. The success of this project also belongs to him and to my daughter, Miriam, who graciously allowed me the time required to complete this work.

To Sue Shaiman, who has been the best mentor I could have imagined. She freely shared her expertise and experience and supported my success at every turn. She allowed me the freedom to follow my own path while still guiding and supporting my progress and supported me with intelligence and empathy. She truly encompasses the definition of 'mentor.'

To Mark Richardson and the Brain Modulation Lab, who gave me the incredible opportunity to research neural data from humans, and who welcomed me, a Speech-Language Pathologist, into the world of neuroscience. Many lab members contributed a tremendous amount of time and effort to my learning, most notably Witek Lipski. Alan Bush and Anna Chrabaszcz, among others, were also phenomenal mentors from whom I learned so much about being a scientist. I am truly grateful.

To my committee and mentors, Robert Turner, Jason Bohland, Bharath Chandrasekaran, Scott Fraundorf and Witek Lipski, I am grateful for your support and feedback which has made this project a success. The process of learning is necessarily confusing and I am grateful for this group of individuals who were always willing to answer questions, no matter how minor.

To my family, without whom I would not be the person I am today. My parents, Angela and LeRoy Dastolfo, have been an incredible support through this process, playing the roles of babysitter, counselor and mentor. They taught me discipline and valued scientific curiosity, which are skills that undoubtably served me in this journey. I am also thankful for my in-laws, Nancy and Bob Hromack, who were more than happy to babysit and offer extra support for our family during this process. Finally, to my siblings, Nick and Maria, who were always there to listen.

To my friends, Yina Quique, Haley Dresang and Caryn Herring, who understood the challenges of a PhD program and always offered a sympathetic ear. I am also grateful for those friends who provided respite from the challenges, Gina Carenbauer, Katie Fish, Lauren Paladino, Rachel Lintelman, Alison Helfrich, Mary Ellen Mitchell-Whisnant, Julie Musarra, Calley Evans and Karen Stickel.

This work was supported through funding received from the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders (NIDCD) F31DC019034 and the Audrey Holland Endowed Student Resource Fund.

Thank you all, from the bottom of my heart.

1.0 Introduction

Prosody, the melody of speech, is an essential component of speech production that permits speakers to express a cornucopia of nuanced intentions; however, the neurological correlates are not well understood. Specifically, it is unknown how neurological processes underlying the prosodic plan (i.e. intensity, intonation, and rhythm) interface with segmental aspects of speech production (i.e. movements related to speech sound production) in order to create the final movements for speech production. Researchers have acknowledged for years that the basal ganglia and motor cortex contribute to the motor control of speech production (Aldridge, Theodoros, Angwin, & Vogel, 2016; Bouchard & Chang, 2014a; D. F. Conant, Bouchard, Leonard, & Chang, 2018; Guenther, 2015; Murdoch, 2001; J. S. Perkell et al., 1997). Speech deficits in patients with Parkinson's disease (Baumann et al., 2018; Jannetts & Lowit, 2014; Pinto, Ghio, Teston, & Viallet, 2010; Watts, 2016) and Huntington's disease (Ludlow, Connor, & Bassich, 1987; Rusz et al., 2013) as well as electrophysiological (Chrabaszcz et al., 2019; Dastolfo-Hromack et al., 2018) and imaging studies of neurotypical and patient populations (Aziz-Zadeh, Sheng, & Gheytanchi, 2010; Golfinopoulos et al., 2015) provide evidence for motor cortex and basal ganglia involvement in prosody. Despite this evidence, our understanding of a specific facet of speech prosody production, phonetic-intensity encoding, is lacking. Phonetic-intensity encoding is a theoretical speech motor control process that describes the *link* between phonetic control and prosodic control (Keating, 2003). Behavioral studies have repeatedly found that as disordered patients and neurotypical individuals increase vocal intensity, articulation improves, demonstrated by changes in acoustic markers of articulation and intelligibility (Dromey, Ramig, & Johnson, 1995; Sapir, Spielman, Ramig, Story, & Fox, 2007; Tjaden & Martel-Sauvageau,

2017; Tjaden & Wilding, 2004), which is behavioral evidence for phonetic-intensity encoding. Speech therapy techniques rely on this phonetic-intensity encoding effect to induce efficient changes across the speech mechanism. Understanding the neurological basis of this effect would 1) provide evidence supporting extant clinical approaches, 2) illuminate the theoretical role of the basal ganglia and motor cortex in speech production and 3) provide brain-behavior correlates for future therapeutic work. This dissertation explores the neurological basis of phonetic-intensity encoding and is composed of a literature review describing neuroanatomy and function of the motor cortex and basal ganglia, as well as theoretical perspectives of prosodic control, followed by an abstract and two experimental studies analyzing the evidence for phonetic-intensity encoding in the basal-ganglia cortical loop and the differential contributions of each region to vocal intensity control.

1.1 Statement of the Problem

Hypokinetic dysarthria affects the vast majority (~89%) of patients with Parkinson's disease (PD)(Logemann, Fisher, Boshes, & Blonsky, 1978; Miller et al., 2007; Schalling, Johansson, & Hartelius, 2017) and causes decreased intelligibility (Barnish et al., 2017; Canter, 1965; Dykstra, Adams, & Jog, 2015) and social isolation (Enderby, 2013; Lirani-Silva, Mourao, & Gobbi, 2015). Given that PD affects approximately 6.1 million people worldwide (Collaborators, 2018), the disease burden is substantial. Although speech therapy has reduced this burden by training a singular therapeutic target (i.e., 'loud' – LSVT®) (Carmichael, Sapienza, & Ramig, 2006; L. O. Ramig, Sapir, Countryman, et al., 2001; L. O. Ramig, Sapir, Fox, & Countryman, 2001), patients in advanced stages and those with deep brain stimulation (Spielman

et al., 2011; Tripoliti, Strong, et al., 2011) do not respond as well to this behavioral intervention, leaving many still suffering from communication deficits. One reason that speech therapy has been effective is because phonatory changes (i.e., increasing intensity) are fundamentally *linked* to increases in articulator movement; therefore, increases in vocal intensity cause widespread improvements in articulation (L. O. Ramig, Sapir, Countryman, et al., 2001; L. O. Ramig, Sapir, Fox, et al., 2001). This *link* has been theoretically described as resulting from the phonetic encoding of prosodic structure (Keating, 2003). However, the neurological basis of phonetic-prosodic encoding remains unknown, leaving a critical gap in fundamental knowledge of speech motor control. The basal ganglia and precentral gyrus are critical nodes in the motor control network and may play a role in phonetic-intensity encoding. Addressing this critical gap would stimulate development of innovative treatments for speech disorders. **The goals of this project are to 1) examine neurological evidence for phonetic-intensity encoding in the precentral gyrus and basal ganglia and 2) test the differential contributions of the basal ganglia (STN) and precentral gyrus for vocal intensity production.** Deciphering the physiology of these speech motor control processes has the potential to not only improve the lives of patients with PD by informing future treatment, but also informing the neurologic basis of speech production.

1.2 Significance of Work

The proposed research will address a critical gap in speech motor control theory, namely, the existence of a process linking vocal intensity control and articulation (i.e. phonetic intensity encoding). Without a complete understanding of this process, therapeutic interventions which currently rely on phonetic-intensity encoding (Boutsen, Park, Dvorak, & Cid, 2018; L. O. Ramig,

1997; Sapir, Ramig, & Fox, 2011) will remain a 'black box' and therefore, limit generalizability of therapeutic effects. Furthermore, speech motor control theory is constructed with the goal of being tested; therefore, the exploration of neurological origins of phonetic-intensity encoding is a natural sequel to the proposition that phonetic-intensity encoding is an observable behavioral process in speech motor control (Keating, 2003). The experiments described in this dissertation are conducted in patients with Parkinson's disease, whose speech deficits are widespread and devastating (Barnish et al., 2017; Canter, 1965; Enderby, 2013; Logemann et al., 1978). Discoveries from this work would not only contribute to speech motor control theories, but also to understanding the neurophysiological processes which underly Parkinsonian speech, informing future therapeutic studies in deep brain stimulation and brain-computer interface technology.

Notable innovative contributions of this work include the evaluation of a theoretical speech motor control process whose existence has been substantiated only through behavioral means (Dromey et al., 1995; Keating, 2003; Tjaden & Wilding, 2004) and inferred from related indirect imaging studies of prosody (Aziz-Zadeh et al., 2010; Guenther, 2015). This work would provide direct electrophysiologic evidence from humans by evaluating vocal intensity control using direct sampling of local field potentials (LFPs) from neuron populations, rather than indirect imaging methods. Vocal intensity control has been evaluated using indirect imaging methods (Aziz-Zadeh et al., 2010; Baum & Pell, 1999; Baumann et al., 2018; Caekebeke, Jennekens-Schinkel, van der Linden, Buruma, & Roos, 1991; Mayer et al., 2002), which do not allow for an analysis of the electrical behavior of neurons. Electrophysiologic recordings (LFPs) offer improved spatial and temporal resolution of neural activity over indirect imaging methods. Analysis of LFPs in this project will provide direct information on neuron population activity, which may be used in future applications (i.e. deep brain stimulation communication outcomes).

1.3 Description of Chapter Structure

The goal of this dissertation was to provide new evidence to support or refute the existence of *phonetic-intensity encoding* and elaborate upon the paucity of research in the neural basis of vocal intensity encoding, which is a component of prosodic encoding. Chapter 2 is an overview of the background and theoretical perspectives underlying prosody production with an emphasis on intensity control and known anatomical substrates for both phonetic and prosodic encoding. The goal of this review is to provide support for the scientific arguments created throughout the course of the study and expand upon theoretical and practical considerations in prosody control, which are integrated in the interpretation of findings. Chapter 3 provides an overview of the study design and aims. Briefly, the study consists of two specific aims, one addressing the primary question of neural correlates for phonetic intensity encoding and a second exploring the differential contributions of the precentral gyrus and subthalamic nucleus to intensity encoding. Chapter 4 is an abstract, which provides a summarized description of study findings and was submitted to the Motor Speech Conference. Chapter 5 outlines an unpublished manuscript arising from the analysis of the primary aim, while Chapter 6 describes an unpublished manuscript centered around the secondary aim, vocal intensity control. The document culminates with Chapter 7, which integrates the knowledge gathered from all study aims and provides final synthesis of study accomplishments and future objectives.

2.0 Background and Theoretical Framework

The following section describes the theoretical and clinical research related to *phonetic-intensity encoding*. The goals are to define the phenomenon as well as describe current interpretations of the phenomenon within common speech motor control theories. Background is also included for the primary decisions made in the study, namely 1) the neurological structures posited to be involved in *phonetic-intensity encoding*, and 2) neurophysiological measurements of gamma and theta band power. Finally, clinical studies of deep brain stimulation are reviewed not only to supplement studies demonstrating that the subthalamic nucleus is important for speech production, but also to investigate the potential clinical relevance of this study.

2.1 Phonetic Intensity Encoding

What exactly is *phonetic-intensity encoding*? It is the *link* (i.e. neural substrate) between neural control of articulatory segments (motoric units related to phonemes), and the suprasegmental features of one component of prosodic control, vocal intensity. Although these two aspects of speech production are often thought of as separate entities, ample evidence exists demonstrating that these processes are intertwined and ultimately planned/programmed simultaneously. The brain must modulate the motor commands that guide the production of phonemic units to account for adjustments necessary to produce the intended prosodic structure. Prosody contains three basic dimensions: pitch, duration and loudness (Hoyte, Brownell, & Wingfield, 2009). For the purpose of this dissertation, this segmentation has been limited to vocal

intensity (loudness) encoding to not only fit constraints of the neural data, but also focus on a prominent feature (i.e. vocal intensity) relevant in the pathophysiology and treatment of speech functions in Parkinson's disease.

Prosody is often described as a suprasegmental feature; however, evidence across multiple studies and domains have shown that this compartmentalization is reductionist. Keating noted that segmental features of speech are modified based on prosodic boundaries (Keating, 2003) and presented evidence from multiple studies of articulatory kinematics demonstrating that phonetic 'strengthening' (i.e. exaggeration or weakening of articulator contact) is dependent upon prosodic boundaries, replicated across four languages (Cho & Keating, 2001; Gordon, 1999; Keating & Shattuck-Hufnagel, 2002; Lavoie, 2001). These claims are supported by older evidence describing that vowel quality changes occur with increased vocal intensity (Lehiste & Universität zu Köln., 1970), as well as recent acoustic studies demonstrating improved articulation (i.e. final word lengthening and intelligibility) with increasing vocal intensity in neurotypical populations (P. J. Watson & Hughes, 2006). The phenomenon of *phonetic-intensity encoding* also aligns with a recent speech motor control observation of 'laryngeal-articulatory coupling' (Dromey, 2010) which reiterates the co-dependence of laryngeal movements (i.e. vocal intensity generator) and articulatory movements in three voice disorders.

Therapeutic interventions have capitalized on this interaction, leading to increases in speech therapy efficiency, particularly for patients with Parkinson's disease. Lee Silverman Voice Therapy (LSVT®) and SpeakOUT® therapy both prioritize increasing vocal intensity as a prominent therapeutic target (Behrman, Cody, Elandary, Flom, & Chitnis, 2020), with LSVT® indicating vocal intensity as the singular target (L. O. Ramig, Fox, & Sapir, 2004). This therapeutic target of 'speak loud' has resulted in improvements in articulation without direct practice in this

domain (Dromey et al., 1995; L. O. Ramig, Sapir, Fox, et al., 2001; Sapir et al., 2002; Sapir et al., 2007). Prior to these approaches, speech therapy targeted the speech sub-systems, structuring therapy goals around either articulation or other prosodic aspects of speech production (Darley, Aronson, & Brown, 1975; L. O. Ramig, Fox, & Sapir, 2008; Trail et al., 2005). Therefore, the *phonetic-intensity encoding* effect has significant relevance to current and future treatment planning.

Neural correlates of *phonetic-intensity encoding* are significantly less documented compared to these behavioral effects. Two studies have suggested that cortico-cortical connections (i.e., corpus callosum) underly the link between linguistic processing, specifically syntactic, and prosody (Klouda, Robin, Graff-Radford, & Cooper, 1988; Sammler, Kotz, Eckstein, Ott, & Friederici, 2010). However, these studies primarily rely on fundamental frequency, which may not display the same neural organization as vocal intensity. Other studies have examined neural correlates of articulation (Brown et al., 2009; Grabski et al., 2012; Hauk, Johnsrude, & Pulvermuller, 2004; Nakayama, Fujii, Suzuki, Kanazawa, & Nakada, 2004; Olthoff, Baudewig, Kruse, & Dechent, 2008; Pulvermuller et al., 2006; Riecker et al., 2000; Takai, Brown, & Liotti, 2010; Terumitsu, Fujii, Suzuki, Kwee, & Nakada, 2006; Wildgruber, Ackermann, & Grodd, 2001) and vocal intensity (Behroozmand et al., 2019; Narayana et al., 2010; Pichon & Kell, 2013) separately without discussion of the overlapping nature of *phonetic-intensity encoding.* None of these studies have evaluated the neural basis of the link between these two processes, which is the focus of the current dissertation. Future sections will discuss the known neural substrates underlying speech motor control in more detail.

As noted, Keating argued that the omission of the phonetic encoding of prosodic structure in theories of speech production (Levelt, Roelofs, & Meyer, 1999) is a fundamental flaw (Keating,

2003); the available data demonstrating the behavioral effects of *phonetic-intensity encoding* are not explained in current theoretical structure. She noted that Levelt (1999) described *phonological encoding* and *phonetic encoding* as two distinct processes, with prosodic planning included in the *phonological encoding*. However, the theory envisions parallel, but separate planning process for prosodic features and articulatory movements, with no explicit link. The articulatory planning is proposed to be retrieved by 'gestural scores' which are stored exemplars of motoric subcomponents of speech sounds. This concept of stored exemplars for speech sounds is recapitulated in more recent models of speech production (Bohland, Bullock, & Guenther, 2010; Guenther, 2015; Perkell, 2013; J. S. Perkell et al., 1997; J. S. Perkell, M. Matthies, M. A. Svirsky, & M. I. Jordan, 1995a) and although the exemplars are modifiable through practice, there is no explicit link between prosody and articulatory changes in established, mature speech systems. To understand the theoretical perspectives of prosodic control more thoroughly, this review discusses prominent and influential theories in speech motor control.

2.2 Theoretical Perspectives

A vast literature exists on prosody and articulatory control including sociolinguistic and cognitive perspectives; however, this review will focus on the goal of understanding *motor control,* and will therefore, limit the scope of the discussion to studies involving three of the most prominent and influential theories of motor control: *Schema Theory, Dynamical Systems,* and the *Internal Model*. The primary goal is to explain articulatory and prosodic movements within a motoric system, not cognitive and linguistic components. These theories will contribute to the interpretation of *phonetic-prosodic encoding* in the basal ganglia cortical loop.

2.2.1 Schema Theory

Speech production, including prosodic and articulatory components, have been interpreted within the framework of *Schema Theory* in numerous studies (Ballard, Robin, McCabe, & McDonald, 2010; Coburn, 2016; Damron, 2004; Meigh, 2017; Sevald, Dell, & Cole, 1995; Welch, 1985). The *Schema Theory* of motor control was introduced by Schmidt and intended to reconcile deficits in *Adam's Theory,* which explained motor behaviors through memory systems (R. Schmidt, 1975). *Adam's Theory* did not match with experimental evidence demonstrating high degrees of variability in learned motor tasks, and it also was not intended to explain ballistic, highly skilled movements, such as speech production. Schmidt's response to these problems was the *generalized motor program* and corresponding *Schema Theory.*

Generalized motor programs (GMP) address the combinatorial explosion problems related to a one-to-one mapping of movement to motor program. GMPs articulate a class, or group of movements, not just one. From this perspective, a range of movements are learned together and categorized as a *motor class*. Schmidt gives the example of "overarm pattern" in throwing a baseball for one potential motor class of movements (R. Schmidt, 1975). This conceptualization would account for the trial-by-trial variability and flexibility that is evident in all skilled motor tasks. However, this requires that the GMP can be modified for different contexts, as the adaptation of the GMP would be necessary.

Parameterization of the GMP allows for this adaptation to occur in novel contexts. Changes in movement speed, or relative force are considered a parameter change to the GMP and are very relevant to the concept of *phonetic-intensity encoding*, which describes alterations to a potential phonetic GMP based on vocal intensity adjustments. These specific changes are modulated within the *motor schema* by which the GMP operates. The motor schema is a memory

function, which accumulates and stores relationships between variables involved in movement. The key considerations are the *initial conditions, response specifications, sensory consequences* and the *outcome* of the movement (Figure 1). All of these factors are stored with the motor schema and allow for flexibility in movement to adapt to multiple situations. However, the detailed predictions of *Schema Theory* focus on motor learning, and less consideration is devoted to parameter selection in novel contexts. For example, one of the primary predictions of the theory is that the strength of the *recognition memory* should increase 'as a function of both knowledge of results during initial practice, and the quality and amount of feedback received on each trial.' (R. Schmidt, 1975).

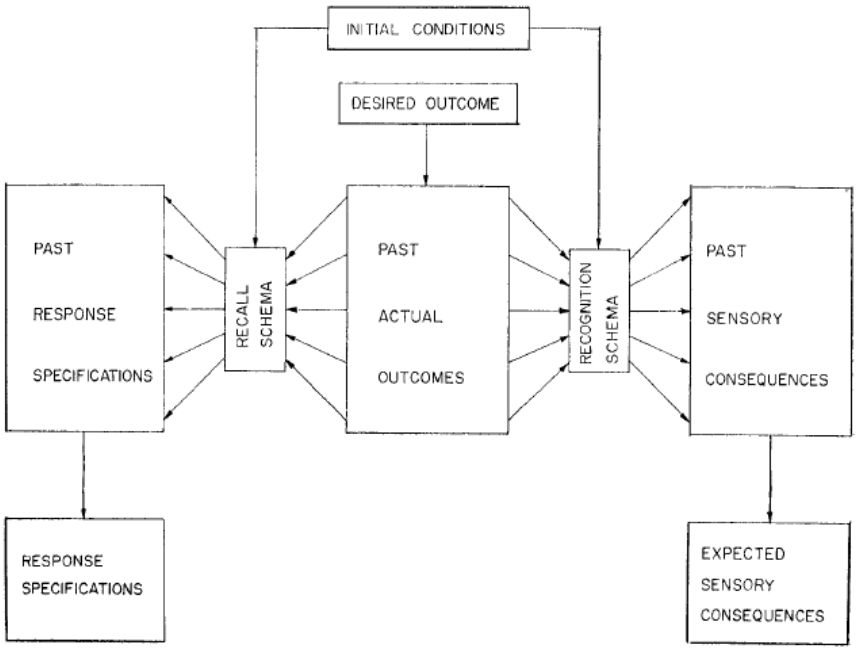


Figure 1. Schema Theory Diagram

Recall and recognition schema diagram. Taken from Schmidt, 1975.

Although schema theory is widely recognized in speech motor control, studies specifically investigating the applications of generalized motor programs in speech prosody have been limited.

Available research explicitly relating *Schema Theory* to prosody has been restricted to syllabic stress (Ballard et al., 2010; Coburn, 2016; Meigh, 2017). Meigh tested the prediction that syllabic stress is contained in the generalized motor program by training participants to produce different stress patterns via a vocal intensity instruction, then testing the *recognition schema* through reaction time tests of auditory word recognition; however, she found co-variance in the reaction time results associated with lists that contained similar phonetic content. It was hypothesized that reaction times would be determined based on syllable stress (old vs. new pattern) since this was the trained parameter; however, reaction times varied based on both stress pattern and phonemic similarity to the training task. Meigh concluded that both syllable stress and phonetic information must have been encoded into the trained GMP, which does not support the conclusion that syllabic stress alone is part of the GMP (Meigh, 2017). She noted that the difference in modality between the training (motor execution) and recognition memory testing (auditory identification task), may have influenced the results. Coburn addressed this concern by replicating Meigh's study, but with transfer testing evaluating the motoric recognition, not auditory recognition. Her results mirrored Meigh's original results, and concluded the same: syllabic stress was encoded with phonetic features (Coburn, 2016).

Importantly, syllabic stress was verified by ensuring that participants were modulating vocal intensity on their trials, providing some support for the concept that *phonetic-intensity* encoding can be interpreted as a united process with articulation within a generalized motor program. Meigh and Coburn both concluded that syllabic stress, at least in isolation, is not part of the motor program and that stress and phonetic similarity are encoded together. According to Schmidt (2003), one of the commonly unaccounted for interpretations of the GMP is the separation between parameterization and generalized motor program. It is possible that syllabic stress, and

therefore vocal intensity, does not represent a component of the GMP, but rather exists as a parameterization of the GMP. In this sense, the *recognition schema* would not encode the specific gradations of intensity associated with each syllable, but rather the phonetic information and *potential* scaling parameters associated with the outcomes. These studies examined learning of non-words. It is possible that the function of prosody is different when the stressed syllable indicates a linguistic change, which cannot be measured in non-words.

Ballard and colleagues conducted a similar style of experiment, training syllabic stress; however, it targeted lexical stress as a potential treatment variable for children with apraxia of speech (Ballard et al., 2010; Coburn, 2016). Theoretically, they make an argument rooted in the *phonetic-intensity encoding* effect, in that prosody affects the entire speech production mechanism, and perhaps training prosody, instead of segmental level errors, may enhance therapeutic gains for apraxia of speech. In Ballard's study, five children underwent syllabic stress training on three syllable non-words, and were later tested using multiple transfer lists, as in Meigh's study. All of the five patients improved in their training and production of the non-words in the stressed-weak syllable condition, leading the authors to conclude that training prosody demonstrated therapeutic efficacy. Some children even transferred some syllable stress patterns (strong-weak) to the real words, which was not expected but highly encouraging.

Examining the results from the perspective of schema theory, some conclusions are apparent. First, syllabic stress (instantiated by vocal intensity) is a trainable variable in the context of motor learning. Second, phonetic content matters. In Ballard's study, transfer occurred best (most accurately) when phonetic content was shared between the trained and transfer stimuli, which replicates the results of Meigh and Coburn. As phonetic content changed to more complex phonemes and real words, transfer degraded. This pattern is consistent with the idea that phonetic

content is somehow encoded with prosodic content in the training process, and by extension, the generalized motor program.

Schmidt reflected on the progress on his theory and acknowledged areas that could and should be improved in subsequent theories of motor control. First, he argued that the concept of a motor program was well supported in the literature. The strongest evidence came from patterning in a limb reach task and observations of the invariant patterning of agonist-antagonist burst that was unaltered by perturbation (Schmidt, 2003). He also argued that invariance can be found in relative timing of GMPs, but not in force. This is important because force is a scaling parameter, and is therefore, a potential correlate of prosody. The difficulty in finding invariance in force may also reflect the difficulty in finding evidence for prosodic stress in the GMP.

2.2.2 Dynamical Systems Theory

Phonetic-prosodic encoding can be interpreted within the framework of Dynamical systems theory (DST). DST emerged in the early 1980s as an alternative approach to motor control using the general motor program (Van Lieshout, 2004). Schema theory failed to recognize the synergistic role of muscles and subsystems in the self-organization of a system. DST conceptualized movement in terms of multiple functional systems working together, like a mass-spring system (Byrd & Saltzman, 2003). In disagreement with schema theory, DST does not require a unified central 'controller' or motor program; instead, it relies on the natural oscillating frequencies of the subsystems to self-organize into movement out of chaos. In a generalized interpretation of DST, the self-organizing systems underlying *phonetic-prosodic encoding* could be the laryngeal system and articulatory system, which are already recognized as working synergistically in a process called laryngeal-articulatory coupling (Dromey, 2010).

According to DST, subcomponents of a system will interact in a predictable, interactive manner; in speech, the subcomponents have traditionally been viewed as different articulators (Kelso & Tuller, 1984). The specific pattern or outcome predicted from any interaction between subsystems is that of periodic motion. Therefore, each articulator (i.e. jaw, lips, vocal folds), can be viewed as a synergistic system within itself, as opposed to multiple individual muscles which require programming. These subsystems are therefore hypothesized to each have an intrinsic resting level oscillation (optimal movement), wherein, the predictable oscillating pattern is considered an *attractor state*. Systems also interact synergistically, and the optimized interaction between systems is called the *attractor point* (Kelso & Tuller, 1984).

In disagreement with the more recently discussed internal model (J. S. Perkell et al., 1997; Perkell et al., 1995a), the DST goals of the speech system are not acoustic, but rather movement trajectories, which can be expressed cyclically (Kelso & Tuller, 1984). Trajectories rely neither on sensory feedback, nor memory-based processing, which represent key concepts in this approach. Invariance in the *relationships between systems*; however, is the postulated driver of the speech motor system in the dynamical approach. Within this paradigm, *control variables* do not constitute a single motor program, rather they are constraints on the system such that these interacting *attractor states* are guided through equifinity to achieve the goal of the task (Kelso & Tuller, 1984).

Control variables have been discussed in many different ways, mostly representing the final goal of the movement (Kelso & Tuller, 1984). Control variables also are sometimes referred to as *order variables* because the variable is an organizing principle, not a central 'controller' as in schema theory or internal model (Tilsen, 2009). Regarding speech prosody, the primary control variable investigated has been *speech rhythm*, focusing on the timing of interaction between

various levels of linguistic structure. Rhythm has been instantiated as predictable changes in syllable, word or phrase duration, cyclic intensity modulations in the speech waveform, and cyclical hesitations or pauses within spontaneous utterances (Barbosa, 2007; Liss, LeGendre, & Lotto, 2010; Tilsen, 2009). Across all studies of speech rhythm, *phase relationships* between segments are important in explaining the dynamics of the prosodic system.

Tilsen has investigated prosody through the lens of DST, through borrowing the concept of *gestures* as an underlying unit of speech production from *Articulatory Phonology* (Browman & Goldstein, 1989; Stalzman & Munhall, 1989). He posits that temporal patterns of phonetic structure, stretching from the phoneme up to phrase level, can be described using oscillatory dynamics (Figure 4). The phrase level constitutes the slowest of the oscillating frequencies, with each sublevel oscillating at twice the frequency of the superordinate level (Tilsen, 2009). It has been noted that this relationship closely resembles the oscillating harmonic frequencies of vocal fold vibration, which is highly relevant to the control of laryngeal dynamics, including vocal intensity. Tilsen hypothesized that each prosodic movement (termed 'pi gesture') was an independent motor control process, which interacts with other levels of gestural processing (Tilsen, 2011) in agreement with the concept of *phonetic-prosodic encoding.* Movement amplitude has also been identified as an important modulator of the system, even destabilizing movements at high amplitude (Haken, Peper, Beek, & Daffertshofer, 1996), which is a consideration for the interpretation of vocal intensity.

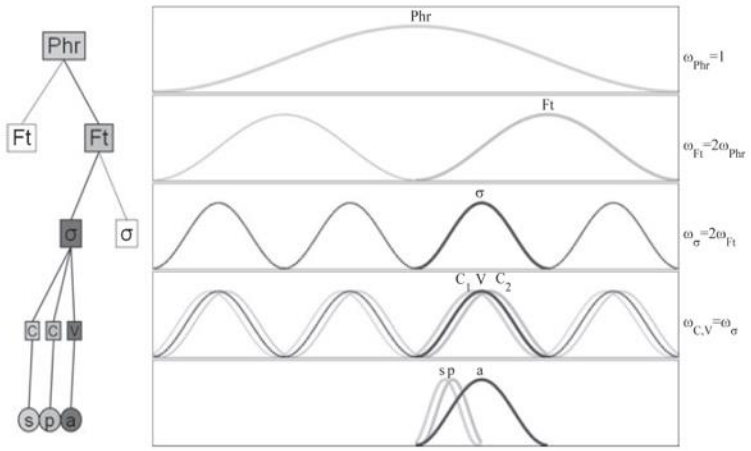


Figure 2. Dynamical Systems Oscillation Example

Relationships between the hierarchical model of prosodic and gestural units. Taken from Tilsen, 2009.

DST has been applied to purely vocal characteristics: register shifts. Register shifts describe the difference between vocal fry, modal, and falsetto vocal production, which are partially dependent upon vocal intensity modulation. Although vocal registers are not incorporated into the hierarchy of Tilsen, it is highly relevant for the production of all vocal modulation, including intensity. Steinhauer analyzed differences in register production from the perspective of registers as *attractor states* (Steinhauer, 2000). Three variables were hypothesized to interact as three attractor states: 1) breathy onset/speech quality/low pitch 2) simultaneous onset/mixed quality/mid F_0 3) breathy onset/falsetto quality/ high F_0. She found that trial accuracy was significantly different based on these combinations of states, but it was dependent upon tasks demands (quality of speech). Therefore, she found evidence that the laryngeal muscles exhibit characteristics of dynamical systems modeling. This study was conducted in non-speech vowel productions;

however, this concept of register transition has not been applied in connected speech, relevant to understanding potential articulatory influences.

2.2.3 Internal Model

The theoretical construct of internal models of movement was introduced in the 1990s, and incorporated both cognitive mechanisms as well as non-cognitive biological constraints (i.e. muscle stiffness) in explaining movement (Kawato & Wolpert, 1998; Wolpert, Ghahramani, & Jordan, 1995; Wolpert, Miall, & Kawato, 1998). Internal models are either *inverse* or *forward models*; forward models have received the most attention for speech research and will therefore be discussed here. *Forward models* of movement dictate that the brain controls movement through ongoing predictions about the 'next state' in movement, and it does so from a stored model which maps sensory information with motor commands (Wolpert et al., 1995). Necessary components of the model to predict, and therefore control, ongoing movement are 1) the current state of the motor system (position of articulators, velocity of movement etc.) 2) the motor command and 3) sensory inflow of information both for ongoing monitoring and updating of the internal model. This internal model concept was introduced to resolve the high computational load of multiple motor programs and was argued to reduce the redundancy in movement and large issues in large degrees of freedom for movement.

In examining internal models, the similarities with schema theory are very apparent. First, the concept that sensory consequences are important in motor learning is reminiscent of *recognition schema* from schema theory. Second, it has been argued that internal models are necessary because reliance on closed loop sensory feedback in ongoing movements is too slow, which has also been argued by Schema theory. Finally, efference motor copy is discussed by both theories as important

to generation of motor commands. Both theories also acknowledge the necessity of long term memory representation of something (motor program or internal model), to explain motor learning in highly skilled movements. Some authors even use the terminology of a *motor program* when working from the internal model perspective (Guenther, 2015). However, differences do exist between these approaches.

Internal models are an even more generalized framework of movement than Schema theory. Schema theory posits that we have generalized motor programs, which specify learned programs for a class of movements; however, internal models do not restrict the learned models to a specific group of movements, it is discussed in an extremely generalized way, not restricting a 'mapping' to a specific context or movement class. Also, despite the fact that Schema theory incorporates sensory information as part of *recognition schema,* internal models emphasize the role of sensation much more strongly. Schema theory uses sensory information primarily in learning and updating the GMP after the movement has occurred; alternatively, internal models treat sensory information as central to the process of modulating movement in an ongoing way, through ongoing state estimation. Internal models incorporate two loops of motor control: 1) an outer loop that uses sensory feedback from external channels, which is similar to the learning and updating which occurs in Schema theory and 2) an internal loop that compares the internal state (movement location) with a sensory prediction of the change due to an ongoing motor command. The second, internal loop allows for online modulation of movement.

Perkell and Guenther have championed the concept of internal models in the realm of speech production (Guenther, 2015; J. Perkell et al., 1997; Perkell, 2013; J. S. Perkell, M. L. Matthies, M. A. Svirsky, & M. I. Jordan, 1995b). One of the main arguments proposed by Perkell is that the goal of speech production is in multi-dimensional auditory-acoustic space (J. Perkell et al., 1997).

This proposal implies that the primary goal of prosodic variation is not movement of muscles, or a certain vocal fold length, but rather fundamental frequency and intensity contrasts, as well as the perceptual correlates of pitch and loudness. He also posits that internal models are required for 'acoustic trajectory' planning, and that both the specifications of the internal model, as well as saturation effects (i.e. when greater movement does not results in a change in perception) determine the final product of speech production. The framework also includes the provision that suprasegmental and segmental features are controlled differently in the speech motor control system (J. Perkell et al., 1997). Perkell uses the example of post-lingually deafened individuals who have received cochlear implants to show that suprasegmental features are more susceptible to sensory deprivation than segmental features.

The most prominent models to emerge from the internal model of speech production are *Directions Into Velocity of the Articulators* (DIVA) and *Gradient Order Directions Into Velocity of the Articulators* (GODIVA) (Bohland et al., 2010; Tourville & Guenther, 2011). The DIVA model primarily focuses on acquisition of segmental features of speech production and is therefore less relevant for the discussion of prosodic features. However, the GODIVA model is intended to explain sequencing of speech sounds, as controlled through the basal ganglia in conjunction with cortical connections. The GODIVA model does not aim to explain prosodic variation, but its highly detailed account of speech production via the looped architecture of the basal ganglia and motor cortex make it a foundational model for studying the basal ganglia and speech production. Because prosody is omitted from the GODIVA descriptions of basal ganglia function in speech, it does not explain how vocal intensity is incorporated into the execution of speech sounds, which was not its original intention. Because of the well-established basal ganglia pathophysiology underlying Parkinson's disease and reduced prosodic variation in the speech of patients as a

predominant characteristic, the basal ganglia are thought to play an important role in prosody, and by extension *phonetic-prosodic encoding*.

Two complementary theories of prosody control have been proposed within the framework of the DIVA model: the *integrated model* and the *independent model* (Guenther, 2015). The neural locus of this prosodic control is proposed to be located within the cortical regions responsible for segmental speech production. Predictions about prosody align with the internal model; a feedforward controller is continually being adapted by somatosensory and auditory feedback error signals. Both models would predict that auditory feedback is imperative to prosodic control, but the independent model separates different channels for pitch, duration and loudness allowing for modulation of one parameter and not the other. The integrated model treats these three parameters as interdependent, summing to a single variable of 'emphasis.' This integrated model would predict that perturbations to one parameter could result in a compensation in a different parameter, as long as emphasis was maintained. Within these models, metrical timing is controlled differently than pitch and loudness variation, because literature has shown that metrical timing is resistant to sensory deprivation (Gammon, Smith, Daniloff, & Kim, 1971).

Different research groups have evaluated the predictions from this model, primarily attempting to differentiate between *integrated control* vs. *independent control* of pitch and intensity (Larson, Sun, & Hain, 2007; R. Patel, Niziolek, Reilly, & Guenther, 2011; R. Patel, Reilly, Archibald, Cai, & Guenther, 2015; S. Patel, Lodhavia, Frankford, Korzyukov, & Larson, 2016) and test intentionality of pitch and intensity control (Hain et al., 2000). These studies do not resolve the question of *integrated vs. independent* control of prosody, but do demonstrate that participants are able to consciously and unconsciously control vocal intensity in ongoing movements, which supports the general tenets of the internal model in prosodic control. Since

prosodic variation, including vocal intensity, is highly sensitive to sensory deprivation whereas segmental features are not, these two components are treated as completely separate systems with the DIVA model. If only considering sensory deprivation effects, this compartmentalization would be appropriate; however, this separation ignores the substantial evidence of co-dependence of prosodic and articulatory performance, and therefore, *phonetic-prosodic encoding.* Therefore, the aspects of these models which attend to vocal intensity control are incomplete.

2.2.4 Summary and Integration

All of these behavioral theories of speech motor control uniquely contribute to the interpretation of *phonetic-intensity encoding*, but none encompass all aspects of this process. For instance, evidence from schema theory shows that articulatory and intensity information may be encoded together (Coburn, 2016; Meigh, 2017), but leaves prosody to 'parameterization,' which is not well described or differentiated from the generalized motor program. The internal model captures the sensory control of pitch and intensity supported through perturbation experiments (R. Patel et al., 2011; R. Patel et al., 2015; S. Patel et al., 2016), but completely omits an interpretation for the known behavioral link between articulatory performance and vocal intensity control. In the DST, a moderate amount of evidence supports the notion that different levels of prosody interact in a cyclical fashion, although none of the studies explained articulatory influences. Furthermore, one of the tenants of DST is that it omits cognition in motor control.

Some aspects of these models will be applied in the current study: 1) the oscillatory nature of DST and 2) the neural locus of vocal intensity control and importance of acoustic information in the internal model. Oscillatory behavior in speech production is relevant to testing neural population effects, which are also proposed to demonstrate oscillatory behavior (Arnal & Giraud,

2012; Cebolla & Cheron, 2019; Giraud & Poeppel, 2012). The proposed study will explore this aspect of oscillatory behavior in more detail, by examining speech through the lens of rhythmic activity contained in local field potentials in the motor cortex and subthalamic nucleus, which will be discussed more in future sections. Furthermore, the importance of acoustic goals (i.e. intensity modulation) and the use of syllables as the fundamental unit of speech production will be utilized in this study as one of the primary variable of interest (vocal intensity), which will allow for applications to these theories of speech production. Results from this study will also inform the core components of these theories. Specific organizational patterns of *phonetic-intensity encoding* may support or refute the proposal of combinatorial control of phonetic information and vocal intensity and provide direction for alterations on these theories based on neuroanatomical correlates.

2.3 Selection of Neurological Structures

2.3.1 Precentral Gyrus Anatomy and Theoretical Importance

Sensorimotor cortex involvement in speech and voice motor control is well established (Bouchard et al., 2016; Chrabaszcz et al., 2019; D. Conant, Bouchard, & Chang, 2014; D. F. Conant et al., 2018), although the precise contributions are yet to be determined. Major neuroanatomical models of speech production recognize the importance of the precentral gyrus in speech production (Guenther, 2015; Jurgens, 2002), and predominantly focus on motor cortex's role in articulatory control. This focus on articulatory control perhaps stems from the original motor control somatotopy identified by Penfield and Boldrey (1937) in which major movement

effectors (larynx, hands, lips, fingers, legs etc.) are represented in the precentral gyrus. Following in this tradition multiple studies have identified articulatory somatotopy in the dorsal precentral gyrus (Bouchard et al., 2016; Eichert, Papp, Mars, & Watkins, 2020; Pulvermuller et al., 2006). More recently, intraoperative studies of speech production have shown that specific phonemes can be decoded from electrophysiologic signals in the precentral gyrus (Chrabaszcz et al., 2019; D. Conant et al., 2014; D. F. Conant et al., 2018), adding to the support that articulatory encoding occurs in the precentral gyrus. In addition to articulatory control, some studies also implicate the left precentral gyrus in prosodic control.

Several neuroimaging and lesion studies have addressed the neurological basis of prosody production (Aziz-Zadeh et al., 2010; Baum & Pell, 1999; Caekebeke et al., 1991; Golfinopoulos et al., 2015; Mayer et al., 2002; Paul, Van Lancker-Sidtis, Schieffer, Dietrich, & Brown, 2003; Pichon & Kell, 2013; Riecker, Wildgruber, Dogil, Grodd, & Ackermann, 2002; Van Lancker Sidtis, Pachana, Cummings, & Sidtis, 2006). Two notable functional magnetic resonance studies found widespread bilateral frontal activation, including the ventral regions of the left motor cortex, during linguistic prosody manipulations (Aziz-Zadeh et al., 2010; Golfinopoulos et al., 2015), which overlaps with areas important for articulation. These studies demonstrate a proof of concept that one potential neural correlate for *phonetic-intensity encoding* could be the overlap between regions specialized for articulation and vocal intensity. It should be noted that these studies focus on emotional and linguistic prosody manipulations, and did not explicitly differentiate between vocal intensity, pitch, and duration. Guenther also noted that the ventral motor cortex (larynx area) may be involved in prosodic control in addition to many of the same regions of interest involved in articulatory control (Guenther, 2015).

Neural correlates of vocal intensity have primarily been investigated through outcome studies in LSVT®. Multiple LSVT outcome studies have found increased activation in bilateral primary motor areas in addition to other cortical speech motor areas following treatment (Baumann et al., 2018; Liotti et al., 2003; Narayana et al., 2010). To disentangle the primary outcome variable in LSVT®, vocal intensity, Narayana and colleagues correlated changes in functional magnetic resonance activation before and after treatment to vocal intensity changes. Correlations were noted bilaterally, but primarily showed a shifted activation toward the right hemisphere and were not apparent in the left precentral gyrus. Although this study suggests that vocal intensity is controlled in the right hemisphere, in agreement with prior literature on generalized prosodic functions (S. Patel et al., 2018; Shapiro & Danly, 1985; Weed & Fusaroli, 2020; Weintraub, Mesulam, & Kramer, 1981; Wildgruber et al., 2005), it fails to define the nuanced effect of *phonetic-intensity encoding*, which is perhaps more related to articulatory organization rather than global prosody. Furthermore, theoretical studies of prosody that focus on the right hemisphere dominant view of prosodic control also acknowledge that prosodic control is unlikely to be completely unilateral and likely involves left hemispheric functions (Heilman, Leon, & Rosenbek, 2004; Wildgruber et al., 2005). Because *phonetic-prosodic encoding* does not simply involve intonational changes, but rather describes changes in articulatory kinematics as a function of prosody, it could be argued that this nuanced effect may appear as subtle modulations in the articulatory organization of the precentral gyrus. This view is supported by one theoretical perspective in prosodic control which dictates that linguistic prosody is modulated in the left hemisphere, while emotional prosody is dominated by the right hemisphere (Heilman et al., 2004; Wildgruber et al., 2005).

Aside from theoretical approaches in prosody, a model derived from both primate vocal control and evidence from humans, also implicates the primary motor cortex. Jurgens formulated a two-

path model for vocal production, one that is dominated by the periaqueductal grey and the limbic system, responsible for motivation and initiation of vocal behavior, and the other pathway devoted to highly skilled fine motor control for speaking (Jurgens, 2002, 2009; Jurgens & Zwirner, 1996). Jurgens highlights that, when cortical motor areas are damaged due to stroke or other neurological injury, complex speech patterns are lost, but patients are still able to produce visceral sounds such as grunting. This neurological presentation is evidence that the cortical motor areas may be important for the highly skilled aspects of speech and vocal control, like *phonetic-prosodic encoding*, which requires an overlapping patterned response combining both phoneme movement and vocal intensity.

Consolidation across experimental studies and theoretical propositions on neuroanatomical correlates of prosodic control would suggest that the precentral gyrus is a logical region in which to investigate *phonetic-intensity encoding.* A plethora of evidence identifies the precentral gyrus as an important modulator of articulatory control, showing clear evidence of the encoding of articulators and phonemes. This articulatory control is the foundation upon which *phonetic-intensity encoding* may exist. Vocal intensity dependent modulations may occur within this articulatory topography, causing the observed phonetic changes that occur with increased vocal intensity. The ventral motor cortex (i.e. precentral gyrus) is also commonly modeled as the final cortical structure in the sequential planning and execution of speech sound production, with output leading to articulatory musculature (Guenther, 2015). If this structure is the final arbiter of motor commands prior to execution, then all information required to produce the precise phonetic structure must be included, including alterations to articulation related to vocal intensity.

2.3.2 Subthalamic Nucleus Anatomy and Theoretical Importance

In addition to the precentral gyrus, the basal ganglia are highly likely to be involved in *phonetic-intensity encoding.* Basal ganglia involvement in prosody production has been well documented due to the effects of lesioning the basal ganglia and speech deficits associated with Parkinson's disease and Huntington's Chorea (Murdoch, 2001). The Jurgens model of vocal motor control, which implicates the motor cortex, also incorporates the basal ganglia, although not the subthalamic nucleus specifically, and it denotes the possibility that the basal ganglia support the generation of the 'correct' vocal sequence (Jurgens, 2009). Furthermore, recent investigations have identified preliminary evidence for articulator encoding in one section of the basal ganglia, the subthalamic nucleus (STN) (Chrabaszcz et al., 2019), including electrophysiological evidence of articulatory gain encoding (Dastolfo-Hromack et al., 2018), one potential instantiation of *phonetic-intensity encoding.*

The STN is a small, subcortical structure, and a functional division of the basal ganglia, which is comprised of the striatum, globus pallidus, substantia nigra and the subthalamic nucleus (Alexander, DeLong, & Strick, 1986; Groenewegen, 2003; Lanciego, Luquin, & Obeso, 2012). The STN is part of the indirect pathway, an important modulator of basal ganglia function (Figure 3), and has recently been identified as a critical member of the hyperdirect pathway (Heida, Marani, & Usunoff, 2008). These pathways allow for communication with the cortex and thalamus (Alexander et al., 1986). Tonic excitation of the STN induces inhibitory output from the globus pallidus, leading to tonic inhibition of the thalamus (Heida et al., 2008). Maintenance of an equilibrium between the *indirect and direct pathways* is hypothesized to underlie motor function, making it an important region of investigation for all motor behaviors, including *phonetic-prosodic encoding*.

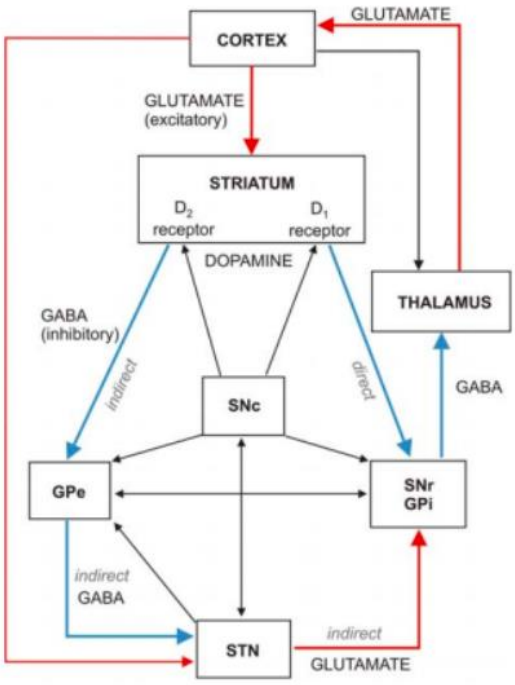


Figure 3. Direct and Indirect Pathways in the Basal Ganglia

Description of indirect and direct pathways, taken from The Subthalamic Nucleus Part II, page 3 (Heida et al., 2008). Excitatory pathways are outlined in red, and inhibitory pathways are outlined in blue. The hyperdirect pathway is designated by the line on the left-hand side connecting the cortex directly to the subthalamic nucleus.

Many hypotheses of basal ganglia function have been proposed (Alexander et al., 1986; Penney & Young, 1983), but some align well with the concept of *phonetic-prosodic encoding*. One of the early hypotheses was *the funneling hypothesis* (Kemp & Powell, 1971), although it has since been invalidated because of the discovery of segregated loops within the basal ganglia circuitry. It was believed that the main function of the basal ganglia was for motor function, and that its role was to gather information from multiple cortical areas (funnel) and select the appropriate context for a specific motor command, which could be interpreted as the prosodic

context in which to modulate commands for phoneme. Other hypotheses focused on the function of the basal ganglia for modulation of movement parameters like speed and amplitude of movement. In the '*scaling and focusing'* hypothesis, it was proposed that the balance between the direct and indirect pathway was responsible for creating the specifications of movement parameters, such as speed. This model may be interpreted as a generalized scaling parameter, in which vocal intensity might be a scaling parameter to the motor commands related to phoneme. The fast direct pathway would activate a motor program, and the indirect pathway would inhibit this action. Critiques of this model stated that the indirect pathway is too slow to balance the fast direct pathway, but with the discovery of the hyperdirect pathway connecting the STN directly with motor cortex, the concept retained legitimacy (M. DeLong & Wichmann, 2009). Finally, the basal ganglia have also been implicated in models of *action selection* (Albin, Young, & Penney, 1989; M. R. DeLong, 1990; Mink, 1996), or more precisely, reduction of possible motor choices (Bar-Gad, Morris, & Bergman, 2003). Again, this interpretation could be applied to *phonetic-intensity encoding* in which inappropriate phonetic motor plans are suppressed to allow for the most appropriately scaled version of the speech movement to be chosen.

These theoretical perspectives converge with electrophysiological, neuroimaging and behavioral evidence of the basal ganglia's role in prosodic control (Aldridge et al., 2016; Dastolfo-Hromack et al., 2018; Liotti et al., 2003; Pichon & Kell, 2013; Rektorova et al., 2012). Neuroimaging evidence has shown that bilateral regions of the basal ganglia are involved in the generation of emotional prosody (Pichon & Kell, 2013), while electrophysiological studies have shown general involvement of the STN in speech production (Chrabaszcz et al., 2019; Lipski et al., 2018; P. Watson & Montgomery, 2006), as well the STN's potential role in modulating phonetic productions measured by the second formant ratio (Dastolfo-Hromack et al., 2018). The

STN is extremely small (approximately 6 mm diameter) making it difficult to investigate through solely indirect methods. The opportunity to directly measure electrical potentials from the STN via awake deep brain stimulation surgery provides a valuable opportunity to expand on these studies and examine the role that this structure plays in *phonetic-prosodic encoding* and speech production as a whole.

2.3.3 Cortical-Basal Ganglia Connections

The proposition of this study is not just that these two regions are working in isolation on different aspects of *phonetic-prosodic encoding*, but rather that they work in concert in a motor control network. The connections of the subthalamic nucleus have distinguished it as a 'hub' of neurological control, having great influence both within the basal ganglia and across broad networks (Baunez, Yelnik, & Mallet, 2011; Bostan, Dum, & Strick, 2010; Jwair, Coulon, & Ruigrok, 2017). The STN has connections to the primary motor cortex, supplementary motor area, premotor area, thalamus, raphe nucleus, pedunculopontine nucleus and cerebellum, in addition to the internal connections of the basal ganglia circuitry (Marani, Heida, Lakke, & Usunoff, 2008).

Cortical motor connections between the STN and the somatosensory area were first identified in the macaque (area 4) (Kunzle, 1976); the result has been replicated, confirming the direct connection between the cortex and STN in primates (Carpenter, Baton, Carleton, & Keller, 1981; Monakow, Akert, & Kunzle, 1978). Studies also have identified more nuanced features to this cortical/STN connection, like somatotopic and topographic organization, which provide the basis of speech motor control models (Bohland et al., 2010). Topographic organization was identified in the primary motor area through anterograde injections into the primary motor area, which demonstrated labeling in the lateral STN (BDA and WG-HRP) (Nambu, Takada, Inase, &

Tokuno, 1996). This topographic organization is replicated throughout the majority of the basal ganglia, such that each motor area representation is somewhat preserved as the information passes through the basal ganglia.

In addition to this topographic organization of the cortical areas, somatotopic organization was found within the STN (Figure 5) and throughout the basal ganglia loops. Researchers found that injection of tracer into three areas (arm, leg, face) in primary motor cortex yielded a preserved somatotopic organization within the STN, and that this occurred both for the supplementary motor area and primary motor area topographic territories (Nambu et al., 1996). Therefore, there is an overlap of topographic and somatotopic organization within the STN. The representation of facial structures is of primary importance to *phonetic-prosodic encoding* as this may underly generation of motor sequences for speech sounds. A recent viral tracing study in the rat STN demonstrated significant contributions to the orofacial regions of primary motor cortex from the rostral-dorsal area of the STN (Degos, Deniau, Le Cam, Mailly, & Maurice, 2008); this projection was found to create a looped architecture with primary motor cortex, just as in the primate model. Additional primate studies include discussion of orofacial movements in the STN, but do not identify the type of orofacial movement elicited from the intracranial stimulation (Nambu et al., 1996; Nambu, Tokuno, Inase, & Takada, 1997); therefore, it is impossible to know if the projections relate to speech musculature (i.e. lips, tongue), or not (i.e. eyelid, eyebrow). This organization provides a potential substrate for overlapping representation of phoneme and vocal intensity control.

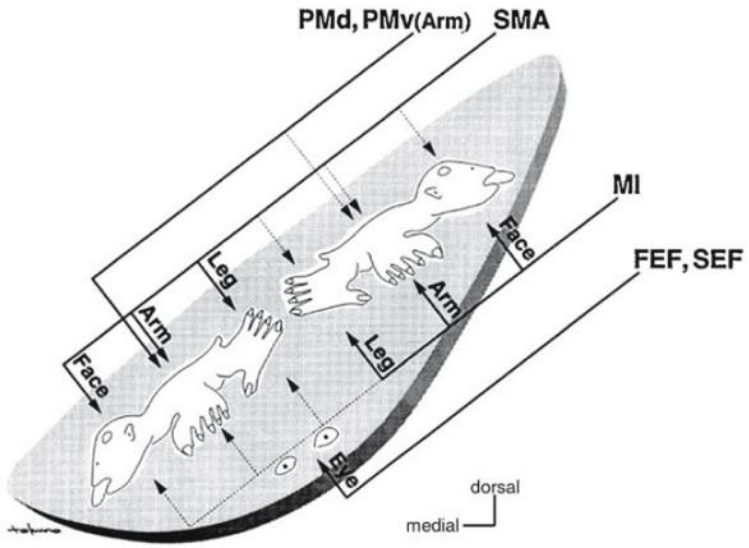


Fig. 16 Mirrored somatotopy in the STN of *Macaca fuscata* (Nambu et al. 1996)

Figure 4. Somatotopy of the Subthalamic Nucleus

STN somatotopy as pictured in Nambu et al 1996, and re-distributed in Marani et al 2008.

Conclusions

Findings from these studies on the structural organization of the STN have propelled theoretical viewpoints of the basal ganglia. Authors have noted that because of the various STN connections, it is uniquely situated to act as a modulator of activity for many neuronal networks (Bostan et al., 2010; Jwair et al., 2017), especially for motor control. The specific functions mediated through these network connections are unknown, but the possibility for speech production and *phonetic-intensity encoding* specifically is high considering the presence of orofacial regions and interconnections with cortical structures. Additionally, the STN connections make control of both limbic and motor functions possible through the motor cortex, frontal lobe, cerebellum, and cingulate cortex connections, which may be involved in prosody generation. Furthermore, the somatotopic designation for oral gestures in the STN-motor cortex affords the ability to integrate multiple sources of information for speech processing. It also raises some

questions regarding the function of these somatotopic areas. How fine-grained is the potential control (i.e. single muscle movements or general oral gestures)? Do these regions apply to speech and voice, or only non-speech gestures (as would be elicited in primates)? The studies that examined oral areas in the somatotopic organization of primates via intracranial stimulation failed to adequately describe the types of oral movements (or sounds) that were elicited. An investigation of the potential role for *phonetic-prosodic encoding* in the STN and motor cortex would provide some evidence to answer these questions of the functional role of these structures in speech production.

2.4 Local Field Potentials and Speech

The basal ganglia-cortical loop has been studied extensively using local field potential recordings, which are the neural measurements used in the present study to evaluate the presence of *phonetic-prosodic encoding.* Local field potentials (LFPs) are electrical signals generated by populations of neurons (Schroeder & Lakatos, 2009). These recordings are of particular importance to deep brain stimulation (DBS) because oscillatory patterns have been shown to relate to different movement parameters (Rosa, Marceglia, Barbieri, & Priori, 2014), and these recordings have also been shown to be biomarkers for adaptive DBS treatment parameters (Hoang, Cassar, Grill, & Turner, 2017). Research using LFPs also provides some direct evidence of neural activity in the subthalamic nucleus, which can inform hypotheses on types of potential movements incorporated in prosodic and articulatory control. LFPs are divided into bands of oscillatory frequencies and studied to understand their role in neuronal communication, as well as the functional and behavioral role a particular frequency may serve (Fries, 2015). Gamma band (70-

150 Hz) and theta band (4-8 Hz) oscillations are the focus of this discussion because these bands have been identified as 'critical' bands for speech production (Giraud & Poeppel, 2012; Sengupta & Nasir, 2016).

Oscillatory activity measured by LFPs is thought to enhance neuronal communication in large-scale networks to modulate multiple brain regions in a synchronous fashion. LFPs reflect net changes in excitability states of neuronal populations, reflected in the summed rhythmic presynaptic change between hyperpolarization and depolarization among a group of neurons (Schroeder & Lakatos, 2009). LFPs have been used to study general neurophysiology in the subthalamic nucleus and motor cortex, as well as attribute specific roles to specific frequency bands. Studies examining LFPs in the STN have attributed changes in oscillation to response inhibition (i.e. release of motor programs), motor vigor, motor effort and motor gating as it is encoded within the subthalamic nucleus (Anzak et al., 2012; Tan, Pogosyan, Anzak, Ashkan, et al., 2013). These concepts, as well as the frequency bands attributed to these processes are highly relevant to *phonetic-intensity encoding* since motor commands for speech may be gated in the same way as limb movements and motor effort or vigor would likely be required to modulate the motor plan for a phoneme.

Much of the LFP discussion in human STN recordings is linked to a process called response inhibition. This is a paradigm during which participants perform motor tasks (such as a button press) after a cue for action (i.e. 'go' response), and then the action is randomly inhibited prior to execution (i.e. 'no go' response)(N. J. Ray et al., 2012). This common paradigm reflects the theoretical role of the STN within the basal ganglia system of movement inhibitor. This paradigm is not well defined for speech production but could reflect the need for sequential motor release signals, as described in the GODIVA model (i.e. "Go Signal"), for speech sound motor programs.

A ‘stopping network’ is thought to facilitate this process which predominantly includes the frontal lobe connections with the basal ganglia (Aron & Poldrack, 2006). Increases in gamma power have been linked to the inhibition (no go) response (N. J. Ray et al., 2012), while other have not seen such an inhibitory gamma response (Brucke et al., 2008). Gamma power is generally expected to increase during movement (Alegre et al., 2013; Anzak et al., 2012; Joundi et al., 2012), but the gamma response is highly variable, evidenced by disparate responses to the inhibition paradigm.

In addition to the response inhibition paradigm, various functional roles have been attributed to frequency bands in the LFP/STN literature. These range in motor functions, but generally focus on the concept that the basal ganglia is responsible for scaling movements (Joundi et al., 2012), in reference to the *scaling and focusing hypothesis* discussed earlier. Movement parameter scaling is a vague concept and has been operationalized in different experimental methods across the literature. Some authors have argued that scaling of movements in the basal ganglia is related not to specific parameters, but rather to overall effort employed to execute the task (Anzak et al., 2012; Joundi et al., 2012; Tan, Pogosyan, Anzak, Foltynie, et al., 2013). It is difficult to understand what ‘effort’ entails and how it should be measured, because it is not clearly delineated in the literature. Anzak discusses the term scaling as ‘best effort of will’ which implies that the movement extent parameters will vary based on movement context (Anzak et al., 2012). Authors have experimentally operationalized ‘effort’ as speed (Joundi et al., 2012), while other operationalize effort as force exerted during movement (Anzak et al., 2012; Tan, Pogosyan, Anzak, Foltynie, et al., 2013). Other authors do not make theoretical connections to ‘effort,’ but study the movement parameters of speed and force on frequency bands independently from the concept of ‘effort’ (Androulidakis et al., 2007; Brucke et al., 2008). Movement scaling is the hypothesized

functional role of the subthalamic nucleus in vocal intensity control and the oscillatory correlates to this process are, therefore, highly relevant to the present study on *phonetic-intensity encoding.*

Conclusions from these STN LFP studies on force, speed and effort have identified some relationships to the frequency band modulations expressed in the STN. When effort is defined as speed, Joundi and colleagues found a significant increase in gamma power related to the onset of an arm-reaching task, and the increase was most prominent for the fast condition compared to normal or slow (Joundi et al., 2012). More recently, a research study examining single unit recordings in the STN found increases in firing related to velocity of a hand reaching task (Potter-Nerger et al., 2017). The relationship between gamma activity and movement speed as well as displacement has also been observed in recordings from the globus pallidus, which receives input from the STN (Brucke et al., 2012).

Additional research papers have found a relationship between frequency bands in STN LFP recordings and force of movement (Anzak et al., 2012; Tan, Pogosyan, Anzak, Foltynie, et al., 2013; Tan et al., 2015; Tan et al., 2016). Anzak et al., who framed the hypothesis in terms of motor effort, found that gamma band activity was related to the overall force in a maximal hand-grip task. They found that high gamma activity predicted both maximal grip force, as well as hand yank force in this task (Anzak et al., 2012). However, Tan and colleagues expounded on their (and others) prior work in force and LFPs relationships in a follow-up study looking at motor effort, combined with a manipulation of effector limb (R vs. L finger). Force parameters generally correlated with LFP response, but the relationship of LFP to effort level was independent of effector of that force, leading the authors to conclude that the STN was encoding effort, not strictly force parameters (Tan et al., 2015).

As evidenced from the previous discussion, gamma band activity has been associated with speed of movement (Joundi et al., 2012) as well as force (Tan, Pogosyan, Anzak, Foltynie, et al., 2013); however, the effects of gamma and movement in the STN are highly variable. The predominant view of gamma and movement in the STN is that it is a 'prokinetic' frequency band and assists with allowing movement to occur (Cassidy et al., 2002; Fischer et al., 2017; Fogelson et al., 2005). This aligns with the commonly held hypothesis that gamma is a 'binding' agent connecting networks related to a task, which has application to both sensory and motor binding for movement execution (Engel, Moll, Fried, & Ojemann, 2005; Engel & Singer, 2001; van Wijk, Beek, & Daffertshofer, 2012). However, some recent evidence demonstrates that increases in gamma band are associated with decreased amplitude of movement (Fischer et al., 2017), which is opposite of the effect one would expect if gamma is prokinetic. Another author proposes that STN gamma band modulation is related to a 'startle response'(N. J. Ray et al., 2012). Therefore, the functional purpose of gamma is still largely unknown and controversial, but it is certainly considered important in the study of movement.

Gamma band activity in the precentral gyrus adheres to a similar pattern compared to the STN. Gamma band power is expected to increase during movement (Ball et al., 2008; Cheyne, Bells, Ferrari, Gaetz, & Bostan, 2008) and is considered prokinetic and related to focal functional anatomy (Brittain & Brown, 2014; Crone et al., 1998; Pfurtscheller, Neuper, & Kalcher, 1993). Pfurtscheller and colleagues found that increases in gamma band power (40 Hz) were topographically distributed over the contralateral precentral gyrus for finger movements, and that the maximal gamma increase occurred prior to movement onset. Importantly, Salari found that movement direction was also encoded in precentral gyrus gamma band (Salari et al., 2019), which

agrees with literature stating that movement fragments, or trajectories can be encoded in single neurons (Hatsopoulos, Xu, & Amit, 2007).

This functional anatomy in gamma band for limb effectors has been replicated for speech articulators across multiple studies (Chartier, Anumanchipalli, Johnson, & Chang, 2018; Chrabaszcz et al., 2019; Salari et al., 2019). Chrabaszcz and colleagues approached this question by contrasting gamma band activity for CVC words which contained 'lip' consonants vs. 'tongue' consonants to identify electrodes which encoded for lips vs. tongue movements (Chrabaszcz et al., 2019), while Chartier and colleagues utilized a decoding method based on kinematic trajectories of articulators (Chartier et al., 2018). The replication across these studies indicates that articulatory representation in gamma band in the precentral gyrus is a robust finding. In addition to articulator encoding, speech acoustics and phonemes have also been identified in gamma band power in the precentral gyrus (Bouchard & Chang, 2014a, 2014b; Bouchard, Mesgarani, Johnson, & Chang, 2013).

Gamma band synchronization is also related to binding processes. Gamma band has been implicated in cortical-STN coupling, with an increase in synchrony between STN single unit firing and gamma band power during movements with fast reaction times (Fischer et al., 2020). It was proposed that this coupling demonstrated evidence of gamma band involvement in movement preparation. Furthermore, these results show that gamma band may be involved in creating the functional motor network between precentral gyrus and the STN. Gamma band is also thought to work in concert with theta band (Canolty et al., 2006; Doesburg, Vinette, Cheung, & Pang, 2012; Landau, Schreyer, van Pelt, & Fries, 2015), another critical oscillation in speech production (Giraud & Poeppel, 2012). Canolty and colleagues found that gamma band power phase-locked to theta oscillations (Canolty et al., 2006). It is possible that these two frequency bands work

together within the basal ganglia cortical loop to modulate numerous aspects of speech production, including *phonetic-intensity encoding.*

Compared to gamma band, research evaluating theta band in motor control is more sparse; however, theta band has been linked to a process called 'motor gating' in the STN. Motor gating is the idea that for any movement to occur, the body must generate a signal indicating that movement has been 'permitted' to occur. Reaction time data has provided primary evidence for this process because the relationship between reaction time, and frequency band modulation would be indicative of the body allowing movement to occur once given a command. Although multiple studies found a significant correlation between the latency of beta desynchronization onset and reaction time (Joundi et al., 2012; Kuhn et al., 2004; N. J. Ray et al., 2012), another found a negative partial correlation of increased theta related to decreased reaction time in a grip force task (Anzak et al., 2012), making this another potential role for theta band. Theta in the STN has also been related to decision-making processes in synchrony with the frontal cortex, including sensorimotor conflict (Cavanagh et al., 2011; Tan et al., 2014; Zavala et al., 2016), 'breaking' once a movement has begun and contribute to the 'hold your horses' aspect of the STN and motor control (Zavala et al., 2013; Zavala, Zaghloul, & Brown, 2015).

Cortical theta band production has been described as 'critical' to speech production. One hypothesis dictates that theta to gamma band synchronization serves to parse the speech signal into decodable units in speech perception (Giraud & Poeppel, 2012) and a similar function is proposed for speech production (Sengupta & Nasir, 2015, 2016). Giraud and Poeppel argue that the rhythmic structure of gamma and theta bands coincides with phrasal and syllabic rhythms of speech perception and note that when phase tracking in theta band does not occur, intelligibility measures sharply decline (Giraud & Poeppel, 2012). Sengupta and Nasir apply these same concepts to

speech production showing that theta and gamma band coherence underlies performance in a speech adaptation task, concluding that these two bands are responsible for feedforward programming (Sengupta & Nasir, 2015). Recent research by the Brain Modulation Lab (the group responsible for overarching NIH study from which these data arise) has implicated both cortical (precentral gyrus) and STN theta band power in a process called articulatory gain encoding (Dastolfo-Hromack et al., 2018; Dastolfo-Hromack et al., 2021), which is the exaggeration of articulator movements associated with clear speech (Sapir, Ramig, Spielman, & Fox, 2010). Collectively, this research highlights the potential influence of theta and gamma band in both the precentral gyrus and STN for *phonetic-intensity encoding.*

Research on neuron spiking data may also inform the process of *phonetic-intensity* encoding in the basal ganglia-cortical loop. Two papers examined single unit activity in the STN during speech production and both research papers documented changes in firing rates during speech production tasks (Lipski et al., 2018; P. Watson & Montgomery, 2006); although the tasks were different (CVC syllables vs. sentence read aloud). Watson and Montgomery reported multiple changes in firing occurring throughout the duration of a sentence, with the hypothesis being that changes may be related to selection of linguistic units and syntax (P. Watson & Montgomery, 2006). Unfortunately, this study was descriptive only and did not employ statistical measures to substantiate the claims. Finally, Lipski and colleagues did employ a statistical approach, separating some of the changes in firing to a relationship with onset of the cue and some with onset of speech (Lipski et al., 2018); increases in firing were observed at the onset of speech production. These results are informative for general STN function in speech production, but may also reflect gamma processes since gamma band activity is thought to reflect single unit firing (Nir et al., 2007).

Conclusions and Critiques

In summary, these neurophysiological studies provide preliminary evidence to support hypotheses that prosodic and articulatory control occur in both gamma and theta band in the precentral gyrus and STN. Prosodic variation involves modulating the speed and force of movements, which has clearly shown a relationship with these bands. Furthermore, studies of LFPs during speech production consistently show an STN response. Cortical data robustly show evidence of articulatory parameter control in the precentral gyrus for speech production, which have been identified largely in gamma band and theta band. Based on this review, prosodic variation may be reflected in gamma and theta frequency bands, which have been shown to be modulated with changes in limb force and speed and capture articulatory control in cortex.

Although these studies are consistent with to the notion that *phonetic-intensity encoding* may be represented in the basal ganglia-cortical loop, further study is required to substantiate this claim as it has not been explicitly tested. In studies on speech production in the STN, stimuli vary from visual presentation and subsequent production of a single letter (Wessel et al., 2016), to no cue and automatic-style speech (i.e. counting months of the year)(Hebb, Darvas, & Miller, 2012), to repetition of a complex sentence (P. Watson & Montgomery, 2006). Stimulus variability makes results difficult to interpret and, by extension, to propose precise hypotheses regarding prosody. Related limb movements studied with LFP analysis often parameterize features of movement, such as force and speech generation (Joundi et al., 2012) because these parameters are hypothesized in theoretical accounts of the basal ganglia function, but the precise correlate of this process in speech production is unknown. Furthermore, studies of motor control in the precentral gyrus have largely focused on articulatory variables. Given that movement direction is a potential control variable in

the cortex (Salari et al., 2019), the direction and extent of phoneme related movements may be modulated in the cortex.

In addition to the variability in speech stimuli and production, few of the studies in the STN test theoretical accounts of speech production. These studies do not reference any specific parameter from speech motor control theories, aside from arguing for the incorporation of the basal ganglia in theories of speech motor control (Lipski et al., 2018). Similarly, the nuances of speech and language processing seem to be ignored based on the stimulus variability including automatic utterances, vs. syllable repetition and sentence repetition. Despite the inconsistencies that exist in the current literature base, evidence supports the notion that the STN participates in speech production, which validates the examination of *phonetic-prosodic encoding* in the STN and precentral gyrus. The current study will expound upon these findings and seek to determine the role of the basal ganglia and the precentral in *phonetic-intensity encoding.*

2.5 Supportive Evidence from Deep Brain Stimulation

Deep brain stimulation (DBS) has had a highly influential and informative impact on the research in basal ganglia involvement in vocal motor control. Patients with Parkinson's disease, essential tremor, and dystonia commonly undergo DBS to treat abnormal motor symptoms through direct stimulation to the subcortical motor circuit. The typical target of stimulation for patients with Parkinson's disease is the STN; therefore, the literature review here is primarily based on STN DBS in Parkinson's disease. DBS has been shown to be very effective in reducing tremor severity and bradykinesia in limb movements in patients with Parkinson's disease; however, the effects on speech symptomology are less clear and highly variable (Aldridge et al., 2016). Pitch

and intensity are important features in the concept of prosody and although *phonetic-intensity encoding* does not explicitly describe pitch, co-variation of pitch and intensity is common in speech (Scharine & McBeath, 2019) and therefore, studies of pitch control may be reflective of larger prosodic processes. The present review of DBS literature will be limited to DBS of the STN and studies discussing these specific parameters in order to inform the current study on both the potential role of the STN in *phonetic-intensity encoding*, but also explore practical applications for the results of this work.

Multiple studies have found positive influences of DBS on prosodic control with regard to pitch variability (D'Alatri et al., 2008; Dromey, Kumar, Lang, & Lozano, 2000). Karlsson and colleagues compared speech outcomes after DBS between STN-DBS vs. cauda zona incerta DBS. They found increased pitch variability in a monologue reading task in STN-DBS, which was independent of linguistic context (i.e. question statements). Patients with targets in the caudal zona incerta did not improve in pitch variability (Karlsson, Olofsson, Blomstedt, Linder, & van Doorn, 2013). Similarly Zhou and colleagues, as well as Gentil et al. tested variability in fundamental frequency and also noted an increase after STN-DBS (Gentil, Chauvin, Pinto, Pollak, & Benabid, 2001; Zhou, Lee, Wang, & Jiang, 2009). Another study found no change in pitch before and after stimulation (Wang, Verhagen Metman, Bakay, Arzbaecher, & Bernard, 2003). The difference between these studies may have been due to low sample size, and therefore low power, in the study by Wang and colleagues (n = 6), compared to the study by Karlsson and colleagues (n = 16). The severity of Parkinson's disease was comparable across studies, and all studies extracted fundamental frequency variability from connected speech tasks, although some differences could also be due to the linguistic factors in stimuli between tasks.

Authors of these studies examining fundamental frequency in patients with Parkinson's disease pre/post STN-DBS have proposed some hypotheses regarding the findings. First, Gentil discusses the possibility that increased fundamental frequency variability after STN-DBS reflects a reduction in rigidity in the laryngeal muscles (Gentil et al., 2001), although they do not provide a definition of rigidity that may be tested in a more rigorous manner. Dromey and colleagues proposed the same explanation for increased pitch variability, but they called it a reduction in 'stiffness" (Dromey et al., 2000). Furthermore, the study by Karlsson and colleagues offers some insight into the anatomical specificity of STN control over pitch variability. The finding that no difference was found in caudal zona incerta DBS, but that increased pitch variability was found in STN DBS, highlights the potentially unique contribution of the STN and pitch variability. However, these studies do not differentiate different types of prosodic variation, and therefore it is difficult to understand how the increase in pitch variability actually contributes to prosody in a meaningful communicative manner (Aldridge et al., 2016).

In addition to pitch, some studies have noted changes in the control of vocal intensity. Wang and colleagues found an overall increase in intensity, measured in a sustained vowel task, suggesting an overall positive effect of STN-DBS on the phonatory/respiratory systems (Wang et al., 2003). However, intensity measured in a sustained vowel context exhibits low ecological validity. A study by Dromey concurs with the results of Wang, but measured changes in intensity from monologue tasks (Dromey et al., 2000). Lundgren and colleagues measured intensity during a reading task pre/post DBS as well as off/on stimulation. They found a modest increase in intensity on/off stimulation (~2 dB), indicating that the stimulator was effecting change; however, there was no difference in intensity between pre-stimulation baseline and long term follow-up one year after testing (Lundgren et al., 2011). Karlsson and colleagues found no significant differences

in intensity on/off STN-DBS stimulation, although a non-significant shallow decline with stimulation was noted (Karlsson et al., 2014). Finally, the Brain Modulation Lab conducted a study evaluating STN DBS outcomes and found electrode lead location did not lead to high perceived intensity, but was related to phonatory airflow, which is a driver of vocal intensity (Jorge et al., 2020). Examining results across studies, the results suggest vocal intensity is impacted by DBS, but no definitive outcome for vocal intensity and STN-DBS can be ascertained.

Theoretically, authors have attributed these changes, or lack of changes, to different processes. Wang and colleagues noted that there may be hemispheric effects to stimulation, with a more appreciable gain in intensity given right sided stimulation alone (Wang et al., 2003), and hemispheric effects were not controlled in the other studies mentioned. In addition to the hemispheric effects, some of these changes could be due to microlesion effects which occur after the DBS electrode is inserted into the brain and (theoretically) disrupts abnormal neural processing. Finally, the effects of levodopa, and other common medications utilized to improve motor symptoms in patients with Parkinson's disease, are also important in interpreting the results. Because of these factors, more research is needed on intensity and STN-DBS.

It is also important to note that prosodic variation is a complex phenomenon, with more aspects than the decomposable characteristics of pitch, duration and intensity. Considering this fact, two studies evaluated prosody by perceptual rating scales. Tripoliti evaluated prosody based on a 1-7 Likert rating scale (7 is normal), and noted a decrease in perceived prosodic variation (by 1 point) at one year after onset of stimulation (Tripoliti et al., 2014). This is informative, but difficult to differentiate effects of stimulation from progression of the disease on prosodic control. This negative effect on prosody is supported by another study which rated prosodic variation on/off stimulation. Santens and colleagues found a perceptually negative effect on prosody, but only

when stimulation was localized to the left hemisphere (Santens, De Letter, Van Borsel, De Reuck, & Caemaert, 2003).

General laryngeal control has also been studied in pre/post STN-DBS, which is important to characterize since prosodic variation is dependent upon phonatory-respiratory physiology. Hammer and colleagues evaluated multiple measures of respiratory physiology in a non-speech task designed to non-invasively elicit subglottic pressure (Hammer, Barlow, Lyons, & Pahwa, 2010). Participants repeated this task both on and off STN-DBS stimulation, and Hammer found that many parameters of aerodynamic speech function changed with stimulation. Participants exhibited greater subglottic pressure, increased peak airflow, and decreased mean vocalic airflow. He concluded that these changes were due to increased respiratory driving pressure from STN stimulation. These modulations in respiratory-phonatory physiology are important pre-cursors for patients to be able to alter pitch and intensity during speech, and therefore could underly *phonetic-prosodic encoding*. Hammer and colleagues noted that the results were consistent with the idea that low frequency stimulation is associated with higher respiratory drive and noted that speech may improve more with low vs. high frequency STN-DBS stimulation. This finding of improved respiratory function following STN DBS was replicated in a study from the Brain Modulation Lab demonstrating increased phonatory airflow pre/post STN DBS (Jorge et al., 2020).

Conclusions and Critiques

In summary, the STN-DBS literature offers some clues to the function of the STN in prosody control, potentially implicating it in *phonetic-intensity encoding*. Laryngeal control is consistently connected to STN DBS, including the primary driver of vocal intensity changes, respiratory function. Research evaluating the role of the STN in vocal intensity control is

inconclusive; however, consistent changes are observed in pitch control with STN DBS which imply at least some involvement of the STN in prosodic functions. Available studies demonstrated low sample sizes and poor controls for articulatory context and stimulation parameters. More research is needed to disentangle the role of STN in prosodic functions.

The largest gap is in this literature is a systematic investigation of different categories of prosodic production, for example linguistic prosody. Multiple studies show changes in pitch variation or intensity modulation with DBS, but the question remains how this intensity modulation is layered onto phonetic aspects of speech production. Furthermore, none of these studies have investigated the possibility of heterogeneity within the Parkinson's population. Disorder subgroups (i.e. akinetic rigid and tremor-dominant) have demonstrated some differential effects in intensity (akinetic rigid subtype showed lower intensity) (Zauber & Huber, 2015), so it is conceivable that some of the variability across outcomes in these different studies is due to heterogeneity within the Parkinson's population.

2.6 Background Summary

Phonetic-intensity encoding is a critical aspect of speech production which is ill-defined in current theoretical accounts of speech production. However, each prominent speech production theory contains components which may afford partial explanations for the co-modulation of articulation and vocal intensity. Schema theory may conceptualize this process as a parameterization of the generalized motor program, while DST may interpret *phonetic-intensity encoding* as the natural oscillation between the complementary systems of articulation and laryngeal control. The internal model may instantiate *phonetic-intensity encoding* as an online

modulation to the internal representation of speech sounds; however, none of these interpretations are described in these perspectives. Substantiation of *phonetic-intensity encoding* through neurological evidence would provide impetus to incorporate it into extant models as well as provide insights for how this process may be instantiated.

Ample evidence exists that the basal ganglia cortical loop is a logical region in which to examine the existence of *phonetic-intensity encoding*. Motor cortex, including the precentral gyrus, has been implicated in fine motor control including numerous accounts of articulatory control, and has been proposed to be important for linguistic prosody manipulations. The STN, and basal ganglia, are thought to be critical for all movements, and are implicated in the modulation of scaling parameters, of which vocal intensity may be a part. Furthermore, numerous studies demonstrate activation in the STN during speech production tasks, as well as articulatory gain, a process which is highly related to *phonetic-intensity encoding.*

This neurological evidence is combined with plausible electrophysiological markers of *phonetic-intensity encoding.* Gamma and theta band have been proposed as two critical bands for speech production. Furthermore, gamma band is the primary measurement linked to articulatory somatotopy in the precentral gyrus, and differentiated articulatory parameters in the STN. Theta band has been implicated in articulatory gain encoding and feedforward programming of speech. *Phonetic-intensity encoding* would likely involve these two bands due to the related findings in the literature. These findings are supported by evidence from deep brain stimulation research that shows both vocal intensity and pitch control are affected in STN-DBS, and that frequency of stimulation differentially affects speech performance.

The experiments described in this document explore *phonetic-intensity encoding* in the STN and precentral gyrus, utilizing gamma and theta band as potential indicators of the shared

neural link between vocal intensity and articulatory encoding. Results will provide relevant knowledge for major theories of speech production and potential speech biomarkers for future studies of deep brain stimulation therapies.

3.0 Study Design

The present study was prospective experimental research designed to test the neurological response in the subthalamic nucleus and precentral gyrus as a function of phoneme and vocal intensity level, which revealed correlates of *phonetic-intensity encoding*. Electrophysiological amplitude is the estimate of neural activity which was applied in a linear mixed effects model as the dependent variable, with phoneme, vocal intensity, and electrode location as predictors. As part of ongoing work in a larger study, local field potential (LFP) recordings in the subthalamic nucleus and sensorimotor cortex were gathered during a speech task in patients with Parkinson's disease during awake surgery for deep brain stimulation implantation. Patients read speech stimuli aloud while electrodes were recording in the sensorimotor cortex and subthalamic nucleus, and patients were prompted to increase vocal intensity on half of the trials, which were balanced for initial phoneme. Recordings were analyzed to examine changes in LFP power as a function of vocal intensity and speech sound -- *electrodes which demonstrated an interaction effect between speech sound and vocal intensity were hypothesized to reflect phonetic-intensity encoding*. It was further hypothesized that the STN will contribute more to vocal intensity modulation compared to the precentral gyrus during speech production. This hypothesis was motivated by the pathophysiology of Parkinson's disease, during which basal ganglia deficits result in reduced vocal intensity output.

Subsequent descriptions detail an experimental investigation designed to empirically determine the functional association between vocal intensity, speech sound and neural population activity. Experimental manipulation of vocal intensity and randomization of articulatory variables provided a causal design by which to conclude the association of theta and gamma band to vocal

intensity encoding and *phonetic-intensity encoding*. Estimations of disease severity were included in models to account for the possibility of pathology driven results. Behavioral speech performance measurements were also included to verify the assumption of improved articulation with vocal intensity increases, as reported in the literature (Dromey et al., 1995; Tjaden & Martel-Sauvageau, 2017).

Electrophysiologic signals of neuronal activity measured directly from the cortical surface and subthalamic nucleus provided information which allows a detailed analysis of spatial topography for *phonetic-intensity encoding*. Electrical activity from neuron populations (i.e., LFPs), are of interest because these signals correlate with speech production (Bouchard & Chang, 2014b; Chartier et al., 2018; Dastolfo-Hromack et al., 2018) and limb movement (Anzak et al., 2012; Tan, Pogosyan, Anzak, Foltynie, et al., 2013). The amplitude of the signal in these oscillatory frequencies is thought to reflect rhythmic, coordinated communication in neuron populations (Rimmele, Gross, Molholm, & Keitel, 2018). LFP amplitude – specifically high gamma band amplitude (termed 'high gamma power') – has revealed differential activation in the primary motor cortex based on speech articulator (Bouchard & Chang, 2014a, 2014b). Differences in gamma power across the primary motor cortex produces 'articulator maps' (i.e., articulator topography) showing that different regions are activated for specific articulators.

Assessment of these variables and measurements from human electrophysiologic recordings allowed for a detailed topographical assessment of *phonetic-intensity* encoding. The mixed effects modeling approach accounted for participant level variability embedded within neural power and described the critical interaction effect between phoneme and intensity. The opportunity to test these hypotheses in direct neurological recordings from patients with Parkinson's disease advances the understanding of *phonetic-intensity encoding*, which is a

fundamental process implicated in numerous clinical interventions (Behrman et al., 2020; Lam & Tjaden, 2013; L. O. Ramig, Sapir, Countryman, et al., 2001).

3.1 Sample Size Justification

To determine the appropriate sample size for the study proposed in the previous section, pilot data were examined to establish statistical power. The resulting information informed the project by testing if the amount of data was adequate to find the proposed three-way interaction between phoneme, vocal intensity and electrode location as described above.

A power analysis was completed using pilot data from 7 participants and 11 sample electrodes. We estimated the effect size for the critical Phoneme x Intensity interaction as partial $\eta^2 = .50$ and the non-sphericity as $e = .80$. A power analysis for a 3-way repeated measures ANOVA using G*Power 3.1.9.3 indicated that the sample size of 29 patients provides more than sufficient power and, in fact, only required 4 participants. A partial analysis of behavioral data showed that vocal intensity production was significantly different between soft and loud conditions in at least 13 patients, with average intensity ranging from 71.55 dB SPL in the soft condition to 73.94 dB SPL in the loud condition (Gates, 2020), demonstrating that the majority of patients modulated intensity during the task. Data were analyzed from all viable data from 29 patients. After elimination of artifactual electrodes, a total of 20 participants with 373 precentral and 111 subthalamic nucleus electrode locations were included in analysis. A thorough description of procedures related to artifact rejection can be found in the methods section and vibration artifact section. Based on the estimation of 4 participants for the critical effect of phoneme by intensity,

this sample size is adequate to answer the questions proposed. Furthermore, similar effects have been found with less data. In pilot data from the Brain Modulation Lab, articulator somatotopy was replicated in the precentral gyrus utilizing 10 participants and 198 electrodes (Chrabaszcz et al., 2019) and articulatory gain encoding was found using the same data (Dastolfo-Hromack et al., 2018). Therefore, it is highly likely that *phonetic-intensity encoding*, which is a derivative of these processes, may also be identified with the available sample size.

3.2 Specific Aims

Once it was determined that the sample size was adequate for the proposed experimental questions, the specific aims described in the study design were expounded upon and described in greater detail. The description of the primary and secondary aims is included in the current section, including detailed descriptions of the independent and dependent variables.

3.2.1 Primary Aim

SA1. Examine neuronal organization of phonetic-intensity encoding of speech production in the precentral gyrus and subthalamic nucleus (Figure 5).

Experimental Question: Do populations of neurons exist in the precentral gyrus and subthalamic nucleus which encode both movements to generate phoneme patterns and vocal intensity increases (*phonetic-intensity encoding*)? We will seek evidence of *phonetic-intensity encoding* by examining LFP power in theta and gamma bands, as these have been identified as critical for speech processing (Giraud & Poeppel, 2012), and have been shown to function in

concert with one another in feedforward adaptation of speech production (Sengupta & Nasir, 2015, 2016).

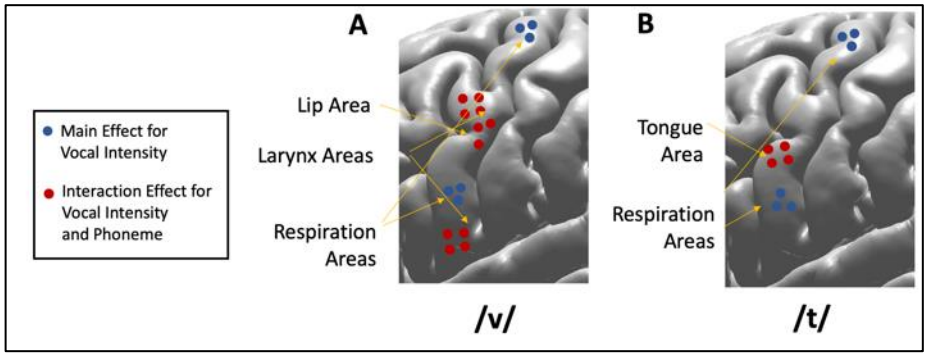


Figure 5. Primary Aim (SA1) Hypothesis

Primary Aim (SA1) Hypothesis: Electrodes in the precentral gyrus will demonstrate phonetic-intensity encoding effects (i.e. interaction effect). LFP power will increase during increases in vocal intensity, but only in articulator regions necessary for a particular phoneme (red dot, interaction effect).

SA1. Statistical Approach.

After raw signal is decomposed into standard frequency components after standard pre-processing, mean-normalized component amplitude (LFP power) will be used to conduct a linear mixed-effects model assessing power changes as a function of 1) phoneme identity 2) vocal intensity and 3) electrode location. A mixed effects model will be fit to account for clustering within the data and a type III analysis of variance will be computed from this model. A recent meta-analysis identified MNI (Montreal Neurological Institute) coordinates of anatomical regions associated with articulators (lips, tongue, respiratory, jaw, larynx)(Guenther, 2015) and these MNI coordinates were originally proposed to be the reference point for articulator coding of electrode positions. The original proposal to determine electrode categorization was based on shortest Euclidean distance from these MNI articulator reference regions. However, these

distinctions yielded coarse labeling with little differentiation between electrodes. A secondary approach was undertaken, utilizing a functional analysis of articulators, which will better account for idiosyncratic differences between participants. A mixed effects model was fit, predicting power in each electrode using known articulators as binary variables (1 = involvement of articulator, 0 = no involvement).

Table 1. Articulator Coding for Functional Analysis of Electrode Locations

Phoneme	Lip	Tongue	Larynx
/s/	0	1	0
/t/	0	1	0
/g/	0	1	1
/v/	1	0	1

$$Neural\ Power = \beta_1[lip] + \beta_2[tongue] + \beta_3[larynx]$$

We hypothesize that articulator region and phoneme variables will show parallel effects in that the electrodes labeled as 'lip region' will increase in power during production of /v/, but not /t/.

Independent Variables (Factors). 1) speech sound /s, t, v, g/ of the experimental trial, 2) vocal intensity on each trial, and 3) electrode location.

Dependent Variable: LFP power

Hypothesis. (Figure 6) Phonetic-intensity encoding will be revealed in the interaction effect between vocal intensity and speech sound. LFP power will increase as a function of both phoneme and vocal intensity, reflecting specific electrodes which encode for movement specification. Furthermore, we hypothesized that this interaction effect of phonetic-intensity encoding will occur

in the precentral gyrus, while the subthalamic nucleus will demonstrate an effect of intensity, but not an interaction with phoneme.

3.2.2 Secondary Aim

SA2. Test the differential contributions of the basal ganglia (STN) and precentral gyrus (PcG) for vocal intensity production.

Experimental Question: What are the respective contributions of the STN and precentral gyrus in the control of vocal intensity output? We will control for PD disease severity using the Unified Parkinson Disease Rating Scale (UPDRS) and time since diagnosis (TSD) because disease severity will likely impact the neural functioning and, therefore, predictive performance in this context.

Equation 1

$$Participant\ Vocal\ Intensity\ (dB\ SPL) = \beta_1(STN\ LFP\ Power) + \beta_2(PcG\ LFP\ Power) + \beta_3(UPDRS) + \beta_4(TSD)$$

SA2. Statistical Approach. Linear Mixed Effect Modeling

Independent Variables. LFP power and disease severity, UPDRS, TSD

Dependent Variables. Participant vocal intensity (root mean square - RMS)

Hypothesis. Neural activation in the precentral gyrus and STN (LFP power) will predict the acoustic intensity level of the motor output, and the STN (basal ganglia) will demonstrate a higher effect estimate reflecting greater contribution. This hypothesis is informed by evidence that the basal ganglia is involved in a speech motor planning loop with the SMA (Chen, Thaler, Nixon, Stern, & Passingham, 1995; Guenther, 2015) which has been implicated in vocal control gain adjustments (Chang, Niziolek, Knight, Nagarajan, & Houde, 2013). It is possible that the

dominant brain region for vocal intensity control will be temporally dependent, with earlier time points favoring the precentral gyrus for motor planning of feedforward commands and later time points favoring the STN for execution of vocal amplitude. For this reason, two models will be conducted in each band of interest, one prior to the onset of speech production and one during speech production.

4.0 Vibration Artifact Rejection

Before testing the hypotheses detailed in the previous section, electrophysiologic data were inspected and processed. Artifact rejection is a normal aspect of processing electrophysiologic data; however, a novel artifact was identified related to the vibration of the electrical system caused by patients' speech which required more extensive work. A working group was established within the Brain Modulation Lab to identify and characterize this novel artifact. The following section describes the artifact and major findings of this working group. Addressing this issue was necessary to ensure the validity of the neural measurement before testing the proposed hypotheses.

The Brain Modulation Lab identified an electrical artifact related to the vibration of the participants' voice production. It was originally observed on a spectrogram and appeared as a high frequency narrow band component, which occurred in the region of the fundamental frequency of the human voice in a three-syllable pattern, identical to the three-syllable triplet pattern in the task stimuli. This narrowband component tracked with the participants' fundamental frequency across trials (Figure 6). In addition to this observation in neural data, the same component was identified in blank recording pins, which were included in some of the participants recording arrays. These recordings were not connected to electrodes, and were therefore not recording any neural data, yet the pattern was still apparent. This three syllable response cannot be accounted for by the frequency following response (Krishnan & Parkinson, 2000; Krishnan, Xu, Gandour, & Cariani, 2004), because it was not observed during stimulus perception and was not localized to the auditory cortex. Finally, vibrations of a similar pattern were intermittently observed after placing a vibration sensor onto the stereotactic frame which holds the participants head in place during

surgery. While the precise mechanisms are still being investigated, the conclusion is that the stereotactic frame and other recording equipment vibrates in response to participants' vocal production therefore contaminating the data. A thorough description of the artifact and analyses used in characterizing the artifact were presented at the Society for Neuroscience Global Connectome 2021 (Bush et al., 2021) and will be submitted for publication.

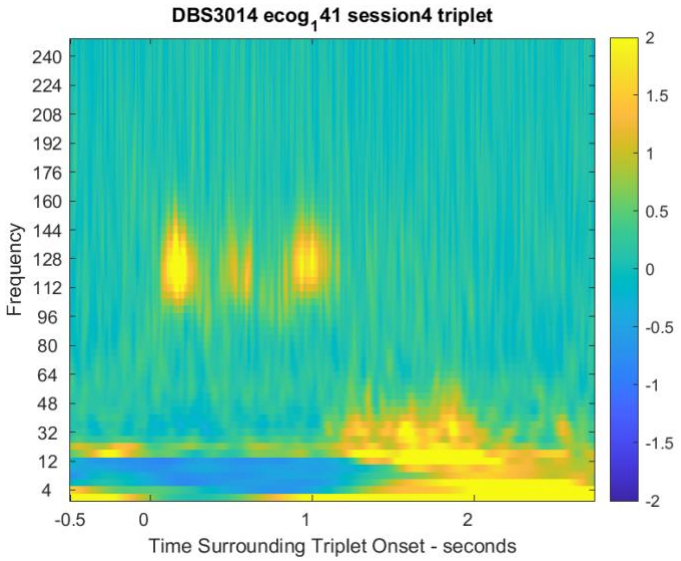


Figure 6. Vibration Artifact Example

Two primary analysis techniques were used to characterize the artifact 1) correlation analysis previously reported by another group who observed the same vibration artifact (Roussel et al., 2020), and 2) a coherence analysis. The correlation analysis examined the cross-correlation matrix between neural power in a broad range of frequency bands to the corresponding power in the audio recording of the participants' voice (Figure 7). The average of the correlation coefficients across the diagonal was used as the artifact contamination score, and permutation testing of this metric was utilized to determine significance of contamination score in a dataset (individual sessions). In addition to this correlation measurement, the coherence between the audio signal and neural signal was calculated for each session. This analysis is systematically

different from the correlation score between it is a direct comparison between signals and is intended to determine causality between each signal, which a correlation analysis cannot accomplish. These two methods were completed on the dataset and dataset measures were significantly correlated to the correlation analysis ($r = 0.55$, $p < 0.001$)(Bush et al., 2021).

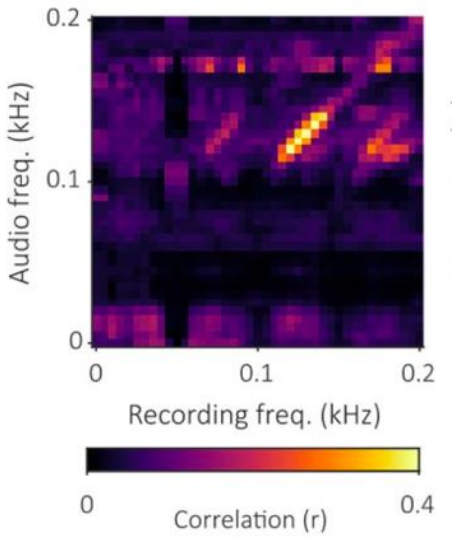


Figure 7. Cross Correlation Matrix

Example cross-correlation matrix taken from figures in (Roussel et al., 2020). High correlations along the diagonal are evidence of vibration related artifact.

To select non-contaminated data for the present study, it was determined that if a true artifact existed, both methods would detect the contamination. The correlation method alone demonstrated a high sensitivity rate (~80%), but a low specificity rate (~20%) when comparing results to the visual analysis of spectrograms as the artifact was originally detected. Combining methods was intended to increase specificity rate. For the coherence method, a significance value was determined to be coherence of 3 based on permutation testing, which correlated to the results of the Roussel method. If a session demonstrated a significant correlation score and coherence greater than 3, it was rejected from analysis. A total of 211 STN recording locations were

originally recorded, with 31 of these locations rejected via the combined coherence and correlation criteria for identification of artifactual influence (14.6% rejected). Out of a total of 631 precentral gyrus electrodes 220 electrodes were rejected for a total of coverage of 410 electrodes represented across precentral gyrus. Brain Modulation Lab scientists conducted a spatial analysis of the electrodes which were identified as artifactual and found no anatomical pattern related to the rejected electrodes.

This amount of data is adequate for the questions presented. The original power analysis, which would have identified the phoneme by intensity interaction effect at a group level, required only four participants. The final number of participants was twenty due to additional artifact rejection parameters (detailed in section 3.1 and methods sections of manuscripts); therefore, this is greater than the four necessary. Furthermore, two previous studies replicated articulator representation on cortex in gamma band (Chrabaszcz et al., 2019) and found evidence of articulatory gain (Dastolfo-Hromack et al., 2018) with only 10 participants and 198 electrodes. These two processes are highly related to the process of *phonetic-intensity encoding* and therefore demonstrate that an effect can be identified in this number of participants and electrodes.

5.0 Abstract

After managing the vibration artifact, a preliminary analysis was completed. This analysis was organized and submitted as an abstract to the 2022 Motor Speech Conference in Charleston, SC. The following section is the submitted abstract for this conference.

Purpose: Speech therapy for patients with Parkinson's disease (PWPD) capitalizes on a physiologic link between vocal intensity and articulation (termed here as 'phonetic-intensity encoding'), causing a patient's articulation to improve given the single objective to speak loudly (Sapir et al., 2007) . However, because specific neurological correlates of this effect remain unknown, phonetic-intensity encoding is not described in models of speech motor control (Guenther, 2015; Hickok & Poeppel, 2015; J. S. Perkell et al., 1997; R. A. Schmidt, 1975) despite its critical theoretical role in prosody (Keating, 2003; Keating & Shattuck-Hufnagel, 2002). The project goals are to 1) determine the neurological organization of phonetic-intensity encoding in the precentral gyrus and subthalamic nucleus (STN), which is critical for movement scaling and 2) explore the differential contributions of the precentral gyrus and STN to generalized vocal intensity control. We hypothesized that phonetic-intensity encoding is represented by vocal intensity-dependent neural activity increases that are topographically specific for particular phonemes.

Methods: Electrophysiologic recordings were obtained from the precentral gyrus and STN in 29 PWPD undergoing deep brain stimulation implantation surgery. During the awake portion of surgery, PWPD repeated at high and low vocal intensity nonword syllables containing the consonants /g, t, v, s/, presented auditorily. To determine if increased vocal intensity improved

patients' articulation, we extracted spectral and temporal measurements - including fricative duration, vowel duration, spectral centroid and the second formant ratio – from the recordings of patient speech, and then correlated them to vocal intensity, measured by syllable-level root mean square. Electrophysiologic signals were decomposed using a wavelet transformation to extract theta band power (4-8 Hz) and gamma band power (70-150 Hz). For Goal 1, linear mixed-effects models were used to predict neural power in each band and brain region from 1) vocal intensity, 2) consonant, 3) recording location and 4) all possible interactions. For Goal 2, averaged neural power by band and brain region were used to predict vocal intensity using a linear mixed-effects model, with two model terms to control for disease severity (time since diagnosis and Unified Parkinson's Disease Rating Scale). To assess for motor preparation and execution effects, the model was fitted to spectral power during the reaction time interval (before speech) and again during speech production.

Results: Increasing intensity improved patients' articulation in both temporal and spectral features. Fricative duration and vocal intensity showed a significant negative correlation (r = -0.16, $p < .001$), and spectral centroid and vocal intensity showed a significant positive correlation ($r = 0.19$, $p < .001$). For Goal 1, the mixed-effects model analysis revealed a significant three-way interaction between phoneme, intensity and electrode location in the STN for theta band power ($F_{(330, 27958)} = 1.39$, $p < .001$), indicating a location-dependent organization of phonetic-intensity encoding (Figure 1). The mixed-effects model in precentral gyrus for gamma band revealed significant two-way interactions for intensity by electrode ($F_{(372, 217840)} = 3.633$, $p < .001$) and phoneme by electrode ($F_{(1115, 218510)} = 4.9309$, $p < .001$), indicating a location-dependent organization of phoneme and vocal intensity encoding in the precentral gyrus. For Goal 2, vocal intensity was significantly predicted by averaged neural power during the reaction interval; neither

region nor band displayed a significant effect. During movement execution, a significant power by band interaction effect ($F_{(1, 200)} = 6.76$, $p < .01$) indicated that gamma band power significantly predicted vocal intensity across both brain regions ($\beta = -0.2$, $p = .02$)(Figure 2).

Discussion: Results demonstrate that a neurological correlate exists for phonetic-intensity encoding: increases in STN theta band power at phoneme-specific locations. Speech therapy techniques which require patients to increase vocal intensity may activate the STN, stimulating articulatory improvement. Precentral gyrus gamma band encodes phonemic and vocal intensity information in non-overlapping locations. Negative association between averaged gamma band and vocal intensity may reflect reduction in competing motor programs with increasing intensity. Predictions of phonetic-intensity encoding were driven by increases in theta band power, which strengthens recent findings that theta band is critical for speech production (Doesburg et al., 2012; Giraud & Poeppel, 2012; Lizarazu, Lallier, & Molinaro, 2019). We conclude that the STN is involved in phonetic-intensity encoding.

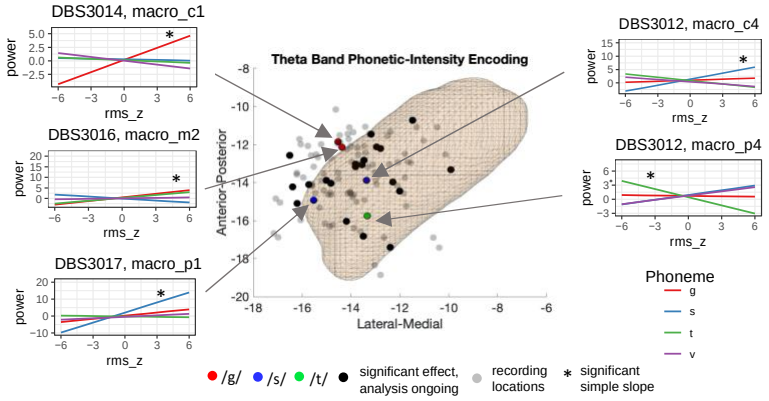


Figure 8. Phonetic-Intensity Encoding Occurs in the Subthalamic Nucleus

Phonetic-Intensity Encoding in the Subthalamic Nucleus (STN). Central plot depicts recording locations in the STN and site at which both phoneme and vocal intensity predicted

neural power (theta band) via an interaction effect. Specific interaction effects for recording locations are depicted in the surrounding figures following simple slope analysis. Raw data points are included as scatter plots for each significant simple slope, color coded to match the slope of interest. Effects were found across participants; participant number is noted by title on surrounding figures.

Figure 9. Gamma Power Negatively Predicts Vocal Intensity Production

Gamma Band Power Negatively Predicts Vocal Intensity Production. A model was fit predicting vocal intensity from theta and gamma band power, averaged across recording locations in the precentral gyrus and subthalamic nucleus. Results showed that gamma band power across both the subthalamic nucleus and precentral gyrus negatively predicted vocal intensity (*). Simple slope for theta band was significant in the precentral gyrus (+), but not when pooled between regions.

6.0 Manuscript #1

Following the submission of the abstract described above, analyses were completed and described in the form of two manuscripts. The first manuscript addresses the primary question of Aim 1, which examines the existence of *phonetic-intensity encoding* within the basal ganglia cortical loop. The second manuscript describes the second aim, which addresses the question of differential contributions of the precentral gyrus and subthalamic nucleus to vocal intensity encoding.

WORKING TITLE: Phonetic-Intensity Encoding Occurs in the Subthalamic Nucleus.

6.1 Introduction

Researchers and clinicians have consistently observed that increasing vocal intensity produces greater articulatory precision in both neurotypical (Dromey & Ramig, 1998; Huber, Stathopoulos, Curione, Ash, & Johnson, 1999) and neurologically disordered individuals (Dromey et al., 1995; Gates, 2020; Maniwa, Jongman, & Wade, 2009; Tjaden, Lam, & Wilding, 2013; Tjaden & Martel-Sauvageau, 2017). This effect relies on a link between the neurological encoding of phonetic and prosodic information (Keating, 2003; Keating & Shattuck-Hufnagel, 2002), which will be referred to as *phonetic-intensity encoding*. Speech therapy interventions, most prominently Lee Silverman Voice Treatment (LSVT®), rely on *phonetic-intensity encoding* to enact highly replicable and meaningful change for millions of people with motor speech disorders by focusing instruction on 'loud speech'(Levy et al., 2020; L. Ramig, Halpern, Spielman, Fox, & Freeman,

2018; L. O. Ramig, Countryman, O'Brien, Hoehn, & Thompson, 1996). Despite the impact of *phonetic-intensity encoding*, the neurological correlates remain unknown. Understanding the neurological underpinnings of this effect would have both clinical and theoretical relevance and enable detailed incorporation of *phonetic-intensity encoding* into theories of speech motor control (Bohland et al., 2010; Guenther, 2015; Levelt, 1992; J. S. Perkell et al., 1997) as well as inform speech therapy intervention. The goals of this study were to 1) investigate the neurological correlates of *phonetic-intensity encoding* within the cortico-basal ganglia loop, specifically the precentral gyrus and subthalamic nucleus (STN), and 2) understand the connection between *phonetic-intensity encoding* and the established articulatory topography of the precentral gyrus (Pulvermuller et al., 2006).

Although the neural correlates of *phonetic-intensity* are unknown, the component parts, phoneme and vocal intensity control, have been well investigated and provide insights for these hypotheses. Prosody has been widely studied through functional magnetic resonance imaging, showing diffuse effects across both cortical hemispheres when producing broad changes in multiple prosodic parameters (Aziz-Zadeh et al., 2010; Baum & Pell, 1999; Caekebeke et al., 1991; Golfinopoulos et al., 2015; Mayer et al., 2002; Paul et al., 2003; Pichon & Kell, 2013; Riecker et al., 2002; Van Lancker Sidtis et al., 2006). However, two studies identified bilateral activation in motor cortical areas, including larynx areas in the left precentral gyrus, during linguistic prosodic manipulation, which implicated regions typically involved in articulation (Aziz-Zadeh et al., 2010; Golfinopoulos et al., 2015). These studies demonstrate the possibility that neurological representations for prosody, namely intensity, may be encoded in the same cortical region as articulatory parameters.

Cortical function has received much attention in prosodic research; however, some researchers propose that the cortex is not the primary driver of prosodic control; rather, the basal ganglia are responsible for prosody generation (Baum & Pell, 1999; Pichon & Kell, 2013). This claim is supported by the monopitch and monoloudness speech characteristics in patients with PD (Liotti et al., 2003), as well as outcomes after pallidotomy (Favre, Burchiel, Taha, & Hammerstad, 2000). Studies conducted in patients with PD echo both these cortical and subcortical claims about prosodic processing, specifically intensity modulation. Outcomes after LSVT® (which utilizes the *phonetic-intensity encoding* effect) demonstrate diffuse cortical changes in neural activity (Narayana et al., 2010; Rektorova, Barrett, Mikl, Rektor, & Paus, 2007) and increased activation in the basal ganglia (Liotti et al., 2003). The basal ganglia are also implicated in vocal intensity because changes in vocal intensity (both increases and decreases) occur following STN deep brain stimulation (Aldridge et al., 2016; Behroozmand et al., 2019). Due to prosodic research implicating cortical motor areas and the basal ganglia, the current study investigated *phonetic-intensity encoding* in the precentral gyrus and subthalamic nucleus.

Studies which evaluate the neural correlates of articulation also support the hypothesis that *phonetic-intensity encoding* may occur within the precentral gyrus (Brown et al., 2009; Grabski et al., 2012; Hauk et al., 2004; Nakayama et al., 2004; Olthoff et al., 2008; Pulvermuller et al., 2006; Riecker et al., 2000; Takai et al., 2010; Terumitsu et al., 2006; Wildgruber et al., 2001). The precentral gyrus, among other motor regions, have been designated as 'regions of high articulatory convergence' based on meta-analyses (Brown, Ingham, Ingham, Laird, & Fox, 2005; Guenther, 2015; Turkeltaub, Eden, Jones, & Zeffiro, 2002), indicating a consensus across studies that these regions contribute to articulator control. Additionally, electrophysiologic (i.e. electrocorticography) studies have identified 'articulator maps' in the sensorimotor cortex,

demonstrating that certain regions of cortex are specialized for specific articulator movements (Bouchard & Chang, 2014a, 2014b; D. Conant et al., 2014; D. F. Conant et al., 2018). Collectively, these studies implicate the precentral gyrus in both prosodic control and articulator movements separately, but questions remain regarding how these regions function to produce simultaneous change in articulation and vocal intensity (i.e. *phonetic-intensity encoding*), which is the focus of the current study.

Few studies address the links between prosody generation and other speech motor control processes. Two studies have proposed that interhemispheric connections (i.e. via the corpus callosum) underly the link between linguistic processing and prosody (Klouda et al., 1988; Sammler et al., 2010). Although it is possible that cortico-cortical connections may also be relevant for *phonetic-intensity encoding*, another plausible hypothesis can be generated from non-speech movement data. Viral tracing studies in primates have shown that the basal ganglia and cortical motor areas are anatomically connected and somatotopically organized in a reciprocal fashion (Alexander, Crutcher, & DeLong, 1990; Alexander et al., 1986). These motor areas (e.g. supplementary motor area, precentral gyrus) have also been implicated in articulatory planning and execution (Bohland et al., 2010). Furthermore, primate data has shown that increased amplitude of movements (which is similar to increased vocal intensity), is achieved through increases in single neuron firing rates in the basal ganglia and greater recruitment of activity (M. R. DeLong, Crutcher, & Georgopoulos, 1983; Georgopoulos, DeLong, & Crutcher, 1983; Panigrahi et al., 2015). It is plausible that the cortical motor areas implicated in articulatory control are modulated by the basal ganglia as a function of vocal intensity.

To investigate this hypothesis, gamma band power and theta band power were selected as primary neural measurements, extracted from local field potential recordings (LFPs). Gamma

band power encodes articulatory information in the precentral gyrus (Bouchard & Chang, 2014a; D. Conant et al., 2014) and is correlated to single neuron unit firing (Nir et al., 2007; S. Ray, Crone, Niebur, Franaszczuk, & Hsiao, 2008). Gamma band activity is also closely linked to BOLD signal reflected in fMRI studies (Nir et al., 2007), which will allow comparison of the results to the numerous studies investigating articulatory and prosodic control. Theta band has equivalent relevance and was correlated to naturalistic differences in articulatory gain (i.e. increases and decreases in the second formant ratio), which is one type of articulatory change that occurs during increases in vocal intensity, in both the precentral gyrus and STN (Dastolfo-Hromack et al., 2021). Both gamma and theta band are described as critical oscillations for speech production (Giraud & Poeppel, 2012). These connections with prior research and extant theories will facilitate the interpretation of results within relevant frameworks and facilitate concept integration into theories of speech motor control.

This experiment utilized intraoperative recordings of speech in patients with PD undergoing deep brain stimulation implantation surgery to investigate the neurological correlates of *phonetic intensity encoding*. This goal was achieved through analyzing gamma and theta band power as a function of phoneme, intensity, and electrode location in the precentral gyrus and subthalamic nucleus. Behavioral speech performance was verified using multiple spectral and temporal measurements of articulatory precision and vocal intensity. Results demonstrate that a neural correlate exists for *phonetic-intensity encoding* and that articulator region contributes to vocal intensity control.

6.2 Methods

6.2.1 Participants.

Data were collected from a total of 29 patients presenting to a neurosurgical clinic for deep brain stimulation surgery. Patients were recruited during clinical visits during which they were undergoing evaluation for deep brain stimulation implantation surgery. Inclusion criteria were: 1) a diagnosis of idiopathic Parkinson's disease, 2) recommendation for surgical treatment by a multidisciplinary movement disorder surgery team, which required thorough neuropsychological evaluation and evaluation of PD symptoms on and off medication, 3) neurosurgical target of the STN, and 4) favorable anatomy free of venous obstructions that would impede electrode contact with the cortical surface. Patients were excluded if they were not candidates for deep brain stimulation based on the above criteria or were unable to perform the task. To ensure informed consent occurred without coercion, study participation meetings and informed consent procedures were held with a research coordinator who was not involved in clinical processes. Patients were assured that the decision to participate would not affect their clinical care in any manner. This research study was approved by the University of Pittsburgh Institutional Review Board (PRO13110420) and all participants completed a written informed consent.

6.2.2 Stimuli

Three syllable 'triplets' (e.g. 'see too gah') were constructed from different combinations of 4 consonants (/g, t, s, v/) and 3 vowels (/i, a, u/). These phonemes were chosen because they consist of articulatory features (involvement of the lips, tongue tip, low tongue/jaw, front tongue, and

larynx) and phonetic state spaces (high, low, front, and back vowels; alveolar, labial, and coronal consonants) which have previously been described in the human motor cortex. Phonemes were presented pseudo-randomly, such that the initial phoneme of the triplets were balanced, and each consonant was presented in the initial position 30 times within a session. Participants were also auditorily prompted (by acoustic example) to increase intensity on half of the trials. The intensity increase was balanced within phoneme condition, so that in a single phoneme condition, 15 were normal intensity and 15 were increased intensity (increase of 20 dB SPL). Actual intensity level of the acoustic stimulus presentation varied based on patient, and audio was adjusted to achieve comfortable hearing level for each participant. The magnitude difference between loud and soft productions was constant.

Table 2. Stimuli Characteristics

Consonant	Alveolar	Labiodental	Velar	Fricative	Voicing
/v/	0	1	0	1	1
/t/	1	0	0	0	0
/s/	1	0	0	1	0
/g/	0	0	1	0	1

6.2.3 Pre-operative Training

While in pre-operative care, patients completed a pure-tone hearing screening at 500Hz, 1kHz, 2kHz, and 4kHz at an intensity level of 25- and 40-dB HL using a calibrated tone generator. Initial hearing screenings were completed by the PI (licensed speech-language pathologist) and then a trained research assistant. Patients were instructed that some of the sounds would be soft, and some are loud, and to use a strong, loud voice when prompted by the stimuli. Patients completed practice using the stimuli and were coached as necessary.

6.2.4 Surgical Procedure and Electrode Placement

Prior to surgery, patients were off medication for 12 hours. As dictated by routine clinical care, patients were anesthetized, a surgical frame was placed around the head for stabilization and burr holes were drilled into the skull at the point of safest entry. After waking the patient from anesthesia, a micro-electrode recording (MER) array, consisting of three electrodes in a fixed triangular orientation was inserted into the brain via the burr hole; three macroelectrode contacts were located superior to the macroelectrodes. LFP recordings were gathered from these macroelectrodes. Trajectory of the MER was guided by pre-operative imaging, and continuous electrophysiologic monitoring by a neurophysiologist. Single cell recordings were used to determine location of the electrode at the superior or inferior aspect of the STN. Recording depths were extrapolated using the changes in depth during placement measured on the MER array, combined with depths associated with STN top and bottom (determined by presence of neuron firing patterns associated with STN). Cortical high density 64 contact electrode strips were also placed through the burr onto the cortical surface, targeted toward the sensorimotor cortex. Trajectory of the cortical electrode strip was planned pre-operatively using MRI images, and a corresponding marker was placed on the patient's scalp during surgery to guide placement. After electrodes were in place, a fluoroscopy image was taken for post-operative localization. Patients were instructed to repeat acoustic stimuli, which were presented through earbuds. Recordings were taken at 2-4 depths within the STN, corresponding to 2-4 sessions depending on patient fatigue and surgical concerns. One session consisted of 120 nonsense syllable 'triplets'.

6.2.5 Intra-operative Recording

Local field potentials were recorded from the deep brain stimulation lead after placement, and this was completed using the Grapevine neural interface processor (Ripple LLC, Salt Lake City, UT, USA). Cortical electrode strips were microphonic-free to reduce movement artifact. Stimuli of triplets were delivered acoustically to the patient through ear buds (ER38-14F foam tips; Etymotic Research, Inc.) after signal was amplified (PreSonus, AudioBox iTwo). Audio recordings were gathered using a high fidelity recorder (H6 Handy Recorder, Zoom Inc., Japan) at a sampling rate of 96kHz, using an omnidirectional microphone in order to reduce distortion related to vocal intensity (MSH-6 Stereo Mic, Zoom Inc., Japan), at 7 cm, 45^{o} angle from the oral angle.

6.2.6 Pre-processing

All electrophysiological data were processed in MATLAB with custom scripts using the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). Data were resampled to a sampling rate of 1 kHz, and cardioballistic effects and line noise were removed using a 1Hz high pass filter and 58-62 Hz notch filter. A high pass Butterworth IIR filter (MATLAB command: 'but') was applied to raw data with a zero-phase forward and reverse filter (MATLAB command: "two pass"), with a filter order of 5. High frequency movement artifacts were identified algorithmically by first calculating the log-normal distribution of the signal amplitude, then determining an appropriate cut off criteria for each participant. The artifact cut off was between 2-3 standard deviations away from the median of the distribution and was determined by examining the distribution and placing the cut off at the point where the tail of the distribution began. Artifact detection was individualized per participant. In addition, a speech production vibration related artifact was

identified, and any session determined to be affected by the vibration related artifact was removed from analysis. A thorough description of the analysis and results of this vibration artifact can be found in Section 4.0. For cortical electrodes, LFP signals were re-referenced offline to a common average reference, applied over blocks of electrodes. Because only three contacts existed for STN recording sessions, a bipolar referencing scheme was applied to the STN contacts. Bipolar references involved averaging the signal with one electrode to the immediately adjacent electrode and subtracting this from the original signal. Data were aligned to the onset of each consonant for epoching. Wavelet transformation was performed using a custom MATLAB function. Wavelet parameters were a width of 7 cycles, evaluated at theta (4-8 Hz) and high gamma frequencies (70-150 Hz)(Lachaux et al., 2007; Nir et al., 2007). Power values were extracted from the wavelet transformation using a custom function (BML lab toolbox)(Bush, 2021) and baseline normalized on a trial-by-trial basis. Power values are derived by calculating the square of the amplitude of the oscillation within the time region of interest, and this was completed within the same function as the wavelet transformation. A baseline period was designated as one second prior to the onset of the stimulus. Power from each syllable segment was normalized to the appropriate baseline for the trial from which it was obtained in order to account for differences in the signal from resting state processes. Boxplots of normalized power values was examined for outliers. Across participants, most outliers appeared above 2.5 standard deviations above the mean (amplitude of 30). A random sample of these outliers were visually examined, and all represented residual movement artifact; therefore, normalized power values above 30 were excluded from analysis.

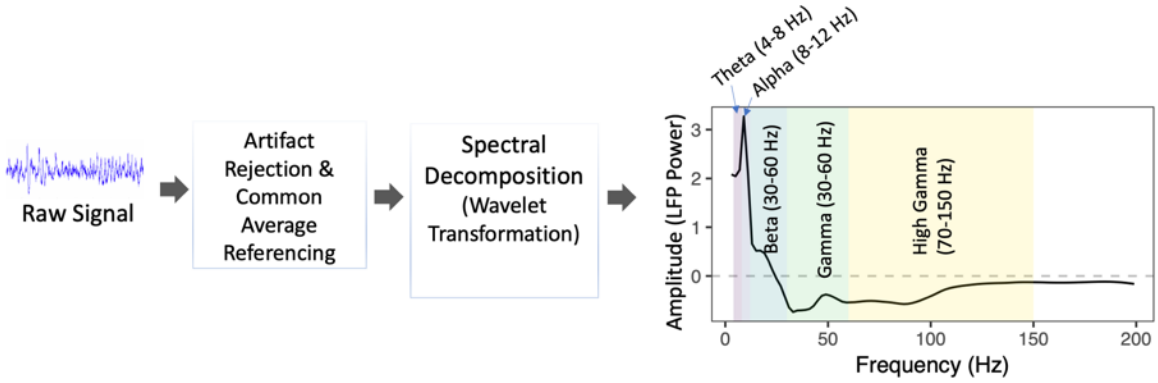


Figure 10. Data Processing Flowchart.

Overview of the processing required to extract the final power values from the raw electrophysiologic signal, which was used as the primary neural measurement.

6.2.7 Electrode Localizations.

Cortical electrode recording locations were determined from reconstructions using intraoperative fluoroscopic images, co-registered preoperative and post-operative computed tomography (CT) images and pre-operative magnetic resonance imaging (MRI) scans, procedures are described in Randazzo et al. (2016). Localization of electrodes into patient native space were conducted using the Freesurfer suite (Dale, Fischl, & Sereno, 1999) (http://surfer.nmr.mgh.harvard.edu/) and a custom-made graphical user interface in MATLAB. All recordings were gathered from the left hemisphere. These localizations were registered into the MNI common space template using Brainstorm (Tadel et al., 2019) (https://neuroimage.usc.edu/brainstorm/). Reconstruction of the subthalamic nucleus recording trajectory was conducted using the Lead DBS toolbox in

MATLAB, (Horn & Kuhn, 2015) based on identification of lead-induced artifact in the post-operative CT scans.

6.2.8 Coding of Speech Data

Phonetic transcription of each participant's produced speech and the markings of speech onset were performed by a trained team of speech pathology students. Using a custom Matlab GUI (github.com/Brain-Modulation-Lab/SpeechCodingApp) (example in Appendix A), onset and offset times were marked based on acoustic evidence of key speech features. Voiced phonemes (i.e. /v/) and vowels were marked at the first visible glottal pulse, indicating the onset of vocal fold vibration. Phonemes that did not rely primarily on voicing were marked based on the characteristic noise features for that phoneme. For example, the rapid increase of high frequency energy denoted the onset of plosive consonants (i.e. /t/ and /g/). A broad band spectrogram was utilized to maximize visualization of key speech features. All coders completed a course in speech science to ensure knowledge of important speech features and were trained in coding procedures by a speech language pathologist. Additional information on coding procedures is available in Appendix A.

6.2.9 Speech Performance Measurement

Vocal intensity level was measured by calculating the root mean square (RMS) per syllable, because prior literature has noted phonetic-intensity encoding effects occur at syllable level (Dromey et al., 1995). Normalized (mean z-scored) RMS was calculated for intensity and assessed within participants. For Equation 2, n = number of samples in the syllable, x = raw amplitude value

at a single sample, μ = mean RMS per participant, σ = standard deviation of RMS values per participant.

Equation 2

$$Equation\ 1:\ RMS\_Z = \frac{\sqrt{\Sigma \frac{x^2}{n}} - \mu}{\sigma}$$

To verify that the *phonetic-intensity encoding* effect similar to what is observed in LSVT® efficacy literature also occurred in the present experimental manipulation, temporal and spectral measurements of speech performance were gathered from the data. Literature evaluating patients with Parkinson's disease and neurotypical individuals showed that with increased intensity, vowel space and vowel duration increase (Dromey et al., 1995; Gates, 2020; Sapir et al., 2010; Tjaden et al., 2013). The current dataset allowed for a trial-by-trial computation of the second formant ratio as a measurement of vowel articulation, therefore this was the chosen measure of vowel performance. Furthermore, improvements in consonant precision also occur with increasing intensity. The two measurements reported in the literature that are also possible to compute within the present dataset are the spectral centroid (Tjaden & Martel-Sauvageau, 2017) and fricative duration (Dromey et al., 1995). Vowel duration was measured from the onset of the glottal pulse following decrease of spectral energy for the preceding consonant until the final detectable glottal pulse indicating the end of vowel duration. Fricative duration was gathered from /v/ and /s/ because these measurements were found to significantly decrease in patients with Parkinson's disease following LSVT® therapy (Dromey et al., 1995) and distinguish clear from conversational speech (Maniwa et al., 2009).

Spectral Centroid

The first spectral moment coefficient (i.e. spectral centroid) was gathered from productions of /t/ and /s/ because the spectral centroid frequency of these two consonants have been shown to increase with increasing vocal intensity in patients with Parkinson's disease (Tjaden & Martel-Sauvageau, 2017). The spectral centroid is calculated as the mean frequency, weighted by the absolute spectrum. It was calculated using a MATLAB script, adapted from a python package to apply to these data (endolith, 2009). Spectral centroid was calculated for /t/ and /s/ because spectral centroid in these phonemes were previously shown to significantly increase with increases in vocal intensity in patients with Parkinson's disease (Tjaden & Martel-Sauvageau, 2017).

Second Formant Ratio

The second formant frequency was extracted from the middle 1/3rd portion of the vowels in each syllable using a custom script in Praat. Random productions were visually inspected for accuracy. The second formant ratio was calculated on a within-subject, trial by trial basis, by dividing the second formant of /i/ by the second formant of /u/. Only trials that included exactly one /i/ production and one /u/ production were used to calculate a trial level F2 Ratio. In order to control for phonetic context, which greatly affects both the first and second formant values (Hillenbrand, Clark, & Nearey, 2001; Strange et al., 2007), the consonants utilized to calculate the ratio were limited to /s/ and /t/. Large amounts of variability in formant frequency is attributable to anterior/posterior place of articulation for initial consonant and voicing patterns (Hillenbrand et al., 2001). Limiting F2 ratio to /s/ and /t/ phonemes maintains a constant place of articulation and consistent de-voicing, limiting the potential influence of phonetic context on formant values. The

ratio was then calculated by dividing the F2 of /i/ by the F2 of /u/ as described in (Sapir et al., 2010).

Outlier Inspection and Rejection

Prior to calculating the articulation metrics outlined above, outliers were removed based on inspection of the boxplot distributions of vocal intensity, measured in root mean square (RMS) across participants. RMS values that exceeded 2.5 standard deviations from the mean were closely inspected for potential noise in the signal. These productions were cross referenced with notations from coders related to noise in the signal. Productions with notations of operation room noise, provider speech, and patient non-speech sounds (i.e. noisy breathing) were excluded from analysis. Of a total 395 values in the outlier range, these criteria excluded 190 productions.

6.2.10 Statistical Analysis

Behavioral Speech Performance

To confirm that patients improved articulation as vocal intensity increased, measures of articulation were correlated to vocal intensity productions. Across all calculations only syllable productions which included the phonemes of interest (/t, s, g, v/) were included for analysis. For each behavioral measurement (fricative duration, F2 Ratio, spectral centroid, vowel duration), a Pearson correlation was computed using the normalized vocal intensity from the full syllable and the measure of interest across all participants to compute one group score. Individual correlations were also computed on a per participant basis and correlation p-values were Bonferroni corrected. Both group level statistics and individual correlations are reported.

Neural Correlates of Phonetic-Intensity Encoding

For the primary goal of investigating the neurological correlates of *phonetic-intensity encoding*, a linear mixed-effects model was fit predicting neural power from 1) vocal intensity (normalized RMS), 2) phoneme identity of the initial consonant, 3) electrode location and 4) all possible interactions. Fixed effects were trial and subject. Assumptions of normality were assessed by visualizing the density distributions and QQ plots of neural power. Electrode location was included in the model to account for the possibility of a topographical influence on *phonetic-intensity encoding*. The original hypothesis stated that the neural effects may depend on electrode location due to the known articulatory topography in the precentral gyrus.

$$
\begin{aligned}
&Neural\ Power \\
&\quad = \gamma_1 Vocal\ Intensity + \gamma_2\ Phoneme + \gamma_3 Electrode \\
&\quad + \gamma_4 Vocal\ Intensity\ x\ Phoneme + \gamma_5 Vocal\ Intensity\ x\ Electrode \\
&\quad + \gamma_6 Phoneme\ x\ Electrode + \gamma_7 Vocal\ Intensity\ x\ Electrode\ x\ Phoneme
\end{aligned}
$$

Following model completion, a type III analysis of variance was conducted to assess overall effect significance for each main effect. The primary effect of interest is γ_7, the three-way interaction effect. A significant three-way interaction effect is considered evidence of *phonetic-intensity encoding*.

Identification of Articulator Locations/Functional Analysis

For the second goal of understanding the connection between *phonetic-intensity encoding* and articulatory topography, a functional analysis was completed to label a subset of electrodes based

on neural response to speech production in the precentral gyrus. Binary variables (i.e. 1 or 0) were created to reflect common articulators identified in the precentral gyrus. Variables were assigned for each syllable based on the articulators required to produce the initial consonant. For example, a syllable with an initial /v/ consonant would have a value of 1 for the lip and larynx variables but 0 for the tongue variable. These variables were used in a linear model predicting neural power in the region and band of interest by three two-way interaction effects between the articulator and electrode location.

$$Neural\ Power = \beta_1[lip\ x\ electrode] + \beta_2[tongue\ x\ electrode] + \beta_3[larynx\ x\ electrode]$$

Electrodes were labeled as a specific articulator if the interaction effect significance was achieved at an alpha level of 0.1 at that specific electrode, and the effect was positive. A positive effect would indicate that an increase in neural power occurred when that articulator was in use compared to the other articulator conditions.

Precentral Articulator Topography and Vocal Intensity

Articulator labeling described above was subsequently used to predict neural power from 1) vocal intensity, 2) phoneme identity, 3) articulator and 4) all possible interactions. Only a subset of electrodes from the precentral gyrus were included in analysis, those which were significantly identified as responding to an articulator. A significant three-way interaction (γ_7) would indicate that *phonetic-intensity encoding* may be driven by articulator encoding.

$$Neural\ Power$$

$$= \gamma_1 Vocal\ Intensity + \gamma_2\ Phoneme + \gamma_3 Articulator$$

$$+ \gamma_4 Vocal\ Intensity\ x\ Phoneme + \gamma_5 Vocal\ Intensity\ x\ Articulator$$

$$+ \gamma_6 Phoneme\ x\ Electrode + \gamma_7 Vocal\ Intensity\ x\ Articulator\ x\ Phoneme$$

6.3 Results

6.3.1 Participants

Intraoperative and acoustic recordings were gathered from 29 patients with Parkinson's disease undergoing STN deep brain stimulation; however, 9 were eliminated due to lack of recordings in the precentral gyrus or based on artifact rejection criteria. Patients were an average age of 65.5 years, with the majority (72%) reporting male gender and right-hand dominance (79%), which is typical of the greater population of patients with Parkinson's disease. Severity of disease was mild-moderate, with an average UPDRS score of 33.7 (SD: 12.6) and average duration of disease reported was 6.5 years. See Tables 4 and 5 for patient demographic information. Supplemental information about the analysis procedures and results for behavioral evaluation are also included in the appendix.

Table 3. Participant Demographics

Participant	Gender	Age	Handedness	Hoehn & Yahr Stage	Duration of disease, years	UPDRS Speech Score	UPDRS Total Score
3001	male	69	right	2	6.5	0	25
3002	male	60	right	NA	5.6	1	29
3003	male	51	left	NA	6.8	1	17
3004	male	61	right	2	10.8	2	61
3005	male	59	right	2	4.8	1	46
3006	female	68	right	2	4.9	1	9
3008	male	70	right	NA	5	NA	NA
3010	male	70	right	NA	1.1	1	34
3011	male	66	right	NA	8.2	1	30.5
3012	male	70	right	NA	20.3	2	43
3014	male	71	unknown	NA	15.4	0	30
3015	male	75	right	NA	6.4	1	42
3016	female	66	right	NA	8.7	2	45
3017	male	76	right	3	0.8	1	34
3018	female	62	right	NA	1.9	1	21
3019	male	71	right	NA	6.9	NA	40
3020	male	64	right	NA	9.3	0	19
3021	female	50	right	NA	4.8	0	6
3022	male	67	right	NA	5.1	1	38
3023	male	61	right	NA	3.7	2	30
3024	female	63	right	NA	5.3	1	37
3025	female	74	unknown	NA	4.5	1	22
3026	male	69	right	2	10.6	1	50
3027	male	57	left	2	3.6	0	33
3028	female	64	left	NA	4.1	2	39
3029	male	78	unknown	NA		NA	NA
3030	female	62	right	NA	8.7	2	43
3031	male	70	right	NA	4.7	2	39
3032	male	56	right	NA	5.8	1	47

Table 4. Patient Speech Performance and Electrode Contacts

Participant	Number of Precentral Electrodes	Number of STN Contacts	Mean Vocal Intensity (raw root mean square)	Speech Latency (seconds)	Three Syllable Duration (seconds)	Number of Sessions
3001	NA	9	0.09 (0.03)	0.63 (0.22)	1.25 (0.21)	4
3002	10	6	0.04 (0.01)	0.44 (0.18)	0.96 (0.10)	3
3003	16	5	0.09 (0.04)	0.90 (0.24)	1.38 (0.14)	4
3004	NA	7	0.18 (0.05)	0.69 (0.26)	1.36 (0.20)	3
3005	NA	NA	0.06 (0.03)	0.58 (0.16)	1.06 (0.17)	3
3006	12	NA	0.21 (0.07)	0.85 (0.23)	1.64 (0.14)	4
3008	7	6	0.08 (0.08)	0.70 (0.21)	1.51 (0.12)	3
3010	35	9	0.02 (0.01)	0.76 (0.33)	1.16 (0.11)	4
3011	16	5	0.07 (0.03)	0.55 (0.17)	0.88 (0.10)	4
3012	17	12	0.15 (0.03)	0.79 (0.18)	1.17 (0.09)	5
3014	6	3	0.15 (0.05)	0.57 (0.33)	1.35 (0.27)	4
3015	7	2	0.05 (0.02)	0.83 (0.23)	1.60 (0.11)	4
3016	10	6	0.13 (0.04)	0.77 (0.57)	1.12 (0.16)	2
3017	24	4	0.26 (0.12)	0.70 (0.67)	1.56 (0.28)	3
3018	17	3	0.15 (0.09)	0.89 (0.20)	1.28 (0.08)	3
3019	15	4	0.05 (0.02)	0.82 (0.25)	1.19 (0.17)	3
3020	49	6	0.01 (0.01)	0.86 (0.35)	1.29 (0.22)	3
3021	22	NA	0.07 (0.02)	0.82 (0.15)	1.29 (0.11)	3
3022	14	6	0.15 (0.09)	0.69 (0.21)	0.98 (0.10)	3
3023	52	NA	0.15 (0.04)	0.62 (0.27)	1.19 (0.25)	3
3024	NA	6	0.14 (0.05)	0.64 (0.11)	1.13 (0.10)	3
3025	NA	NA	0.11 (0.04)	0.66 (0.28)	1.27 (0.29)	3
3026	NA	NA	0.10 (0.03)	1.15 (0.45)	1.23 (0.13)	2
3027	NA	NA	0.06 (0.03)	0.72 (0.32)	1.32 (0.25)	3
3028	NA	NA	0.08 (0.03)	0.66 (0.13)	1.43 (0.09)	4
3029	NA	NA	0.16 (0.06)	0.63 (0.17)	1.38 (0.14)	2
3030	23	3	0.07 (0.02)	0.72 (0.15)	2.12 (0.58)	2
3031	NA	NA	0.18 (0.14)	0.56 (0.17)	1.3 (0.12)	2
3032	16	9	0.14 (0.06)	1.15 (0.46)	1.49 (0.12)	3

6.3.2 Greater Articulatory Precision with Increasing Vocal Intensity

One purpose of the study was to determine if participants adjusted articulatory precision as vocal intensity production increased, consistent with the previously documented effects related to *phonetic-intensity encoding* (Dromey et al., 1995; L. O. Ramig, Sapir, Fox, et al., 2001; Sapir et al., 2007). Vocal intensity data were examined for contamination by background noise and non-speech patient noises, and only clean data were included for analysis; data elimination led to less than 1% data exclusion. Participants produced an average vocal intensity of 0.11 RMS with an average range of 0.26 RMS (Table 5). Vocal intensity productions were correlated to six total spectral and temporal measurements of articulatory precision. A Bonferroni adjusted alpha of 0.008 was utilized to determine significance of the correlated variables.

Temporal Measures of Articulatory Precision

Vocal intensity productions from all participants were combined and correlated to the temporal articulatory variables of fricative duration, applied to the two fricatives in the dataset (/v/ and /s/), and vowel duration (Figure 11). Fricative duration was negatively correlated to vocal intensity for both /v/ duration (r = -.16, p < .001) and /s/ duration (r = -.046, p = .00043). Fricative duration correlation patterns were consistent within individual participants (Figure 11 B&D). Vowel duration was positively correlated to vocal intensity (r = .13, p < .001). The vowel duration correlation was examined within individual vowels. The significant positive effect was present in /a, ə, i/ and a positive but non-significant trend was present in /^/ and /a/ (Appendix B). Vowel duration correlation patterns were consistent within individual participants (Figure 11 F). Outlier influence was examined by replicating correlations after removing participants containing outliers; removing outliers did not change the effect size or significance of any effects. All significant

temporal correlations are consistent with behavioral patterns previously documented in patients with Parkinson's disease, although smaller in effect size.

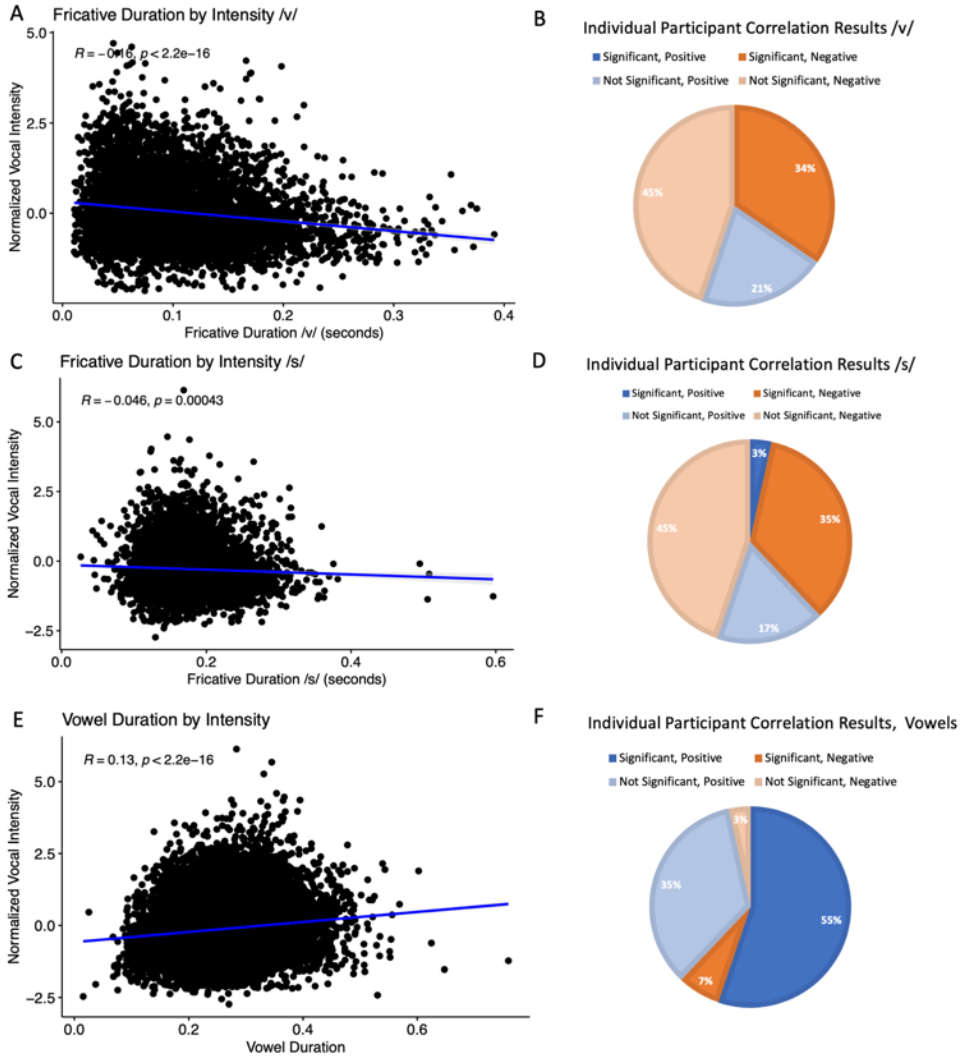


Figure 11. Temporal articulation adjustments with increasing vocal intensity.

A. and C. Correlation between normalized vocal intensity and phoneme duration for the /v/ phoneme (A) and /s/ phoneme (C) and vowel for all participants. B. and D. Percentage of individual participant correlations which are significant after Bonferroni correction and the direction of the correlation effect estimate.

Spectral Measures of Articulatory Precision

Vocal intensity productions from all participants were combined and correlated to the spectral articulatory variables of spectral centroid, applied to /t/ and /s/, and the second formant ratio (F2 Ratio (Figure 12). Spectral centroid was positively correlated to vocal intensity for both /s/ (r = .19, p < .001) and /t/ (r = .035, p = .0041). Spectral centroid correlation patterns were consistent within individual participants (Figure 12 B&D). F2 Ratio was positively correlated to vocal intensity (r = .10, p < .001). The F2 Ratio correlation patterns were consistent within individual participants (Figure 12 F). Outlier influence was examined by replicating correlations after removing participants containing outliers; removing outliers did not change the effect size or significance of any effects. All significant spectral correlations are consistent with behavioral patterns previously documented in patients with Parkinson's disease. Spectral centroid increases are associated with forward tongue motion (Kent & Read, 2002), and the F2 Ratio is associated with greater tongue movements creating distinctions between different vowels (Sapir et al., 2010).

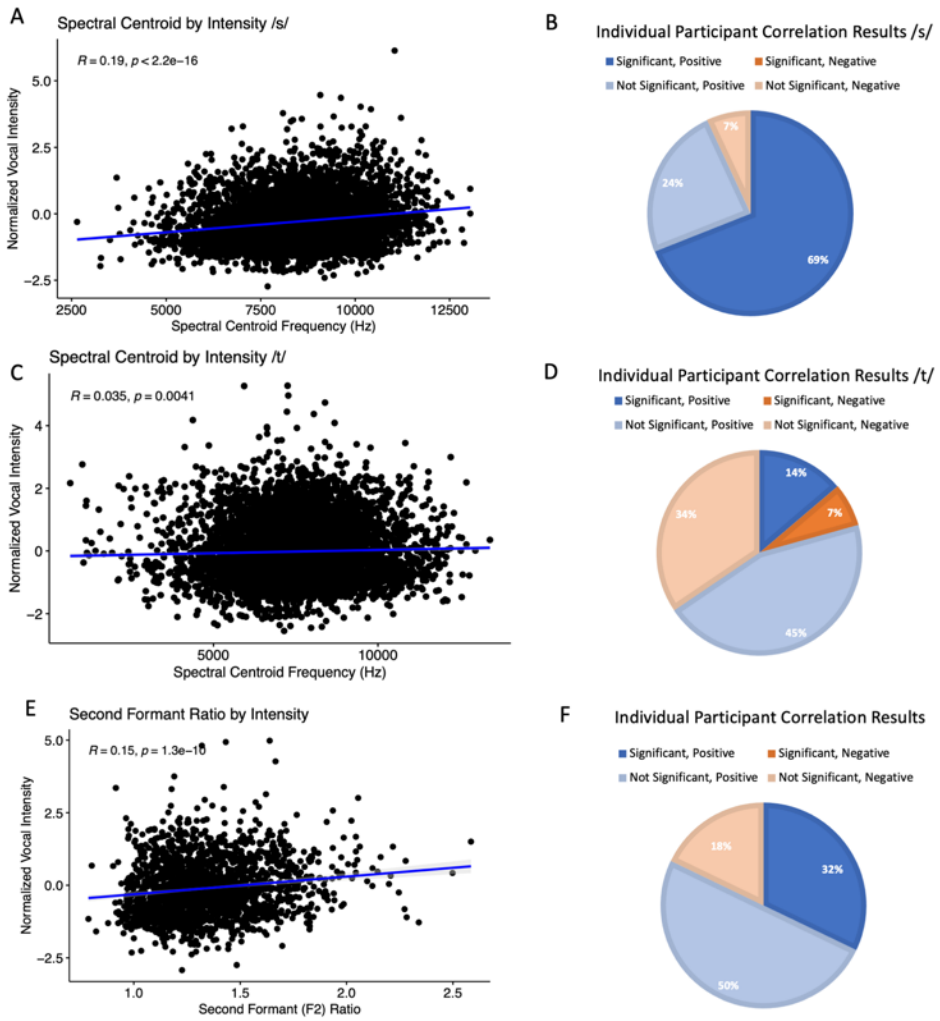


Figure 12 Spectral centroid frequency increases with vocal intensity.

A. and C. Correlation between normalized vocal intensity and spectral centroid frequency for the /s/ phoneme (A) and /t/ phoneme (C) for all participants. E. Correlation between the F2 Ratio and vocal intensity for all participants. B., D. & F. Percentage of individual participant correlations which are significant after Bonferroni correction and the direction of the correlation effect estimate, corresponding to the adjacent measure.

6.3.3 Phonetic Intensity Encoding Occurs in STN Theta Band

To identify a neural correlate of *phonetic-intensity encoding*, it was hypothesized that *phonetic-intensity encoding* would be represented by dual-functioning regions controlling both phoneme and intensity. Neural recording locations which demonstrated vocal intensity-dependent increases in neural power, but only for particular phonemes constituted evidence of *phonetic-intensity encoding*. This effect is expressed as a three-way interaction between vocal intensity, phoneme and recording location. A mixed-effects model was fitted predicting neural power for each band (theta and gamma) and neural region (precentral gyrus and STN), followed by a type III ANOVA to examine overall effects. A significant three-way interaction was found in the model predicting theta band power in the STN ($F_{(330, 27958)} = 1.39$, $p < .001$) (Figure 13 A). This three-way interaction effect was not significant in the other models fit for the precentral gyrus and STN gamma band power. Neural distribution was positively skewed (Appendix B). To ensure that results were not due to the large sample size, the likelihood that randomly sampled variables would produce a significant result was tested. The labels for electrode location were randomized and the model was refit with these parameters for a single iteration. No effects were significant based on these results.

To understand *phonetic-intensity* interaction effects at each of the 24 significant recording locations, follow-up testing was conducted using a simple slope analysis which applies the same model to recording location level data (Figure 13 B-F). The original model converged in only 5 of the significant recording locations. All viable recording location models demonstrated a significant slope for a single phoneme which predicted power differently compared to the remaining phonemes. Two recording sites demonstrated a significant positive slope for /g/ (Slope

Estimate = 0.744, p = .007; Slope Estimate = 0.57, p = .018) and two recording sites demonstrated a significant positive slope for /s/ (Slope Estimate = 1.98, p = .01; Slope Estimate = 0.749, p = .015), indicating that neural power increased with vocal intensity, but only for a single phoneme at these recording locations. One remaining recording location demonstrated a significant negative simple slope for the /t/ phoneme (Slope Estimate = -0.579, p = 0.048), indicating that predicted neural power decreased as vocal intensity increased, but only for /t/. To determine the driving relationships at the recording sites for which the mixed-effects model would not converge, Pearson correlations examined the relationship between power and vocal intensity, separated by phoneme. Correlation results demonstrated that a single phoneme exhibited a significant relationship with power in with 5 of the remaining 19 electrode sites (Appendix B). Positive relationships were observed in 3 sites for /v, g/ phonemes, and negative relationships were observed in 2 sites for /t, s/ phonemes. Electrode sites which did not yield a significant correlation for any phoneme still demonstrated the pattern of a single phoneme showing a stronger relationship with power; marginally significant ($p > 0.05$ and < 0.1) sites yielded the same pattern in an additional 4 sites. Two additional marginally significant sites demonstrate a dual phoneme (/s, v/ and /g, t/) positive relationship with power. Correlation figures are included in Appendix B. All together, these results show that 14/24 electrode sites demonstrated the pattern of a single phoneme relationship with power and intensity.

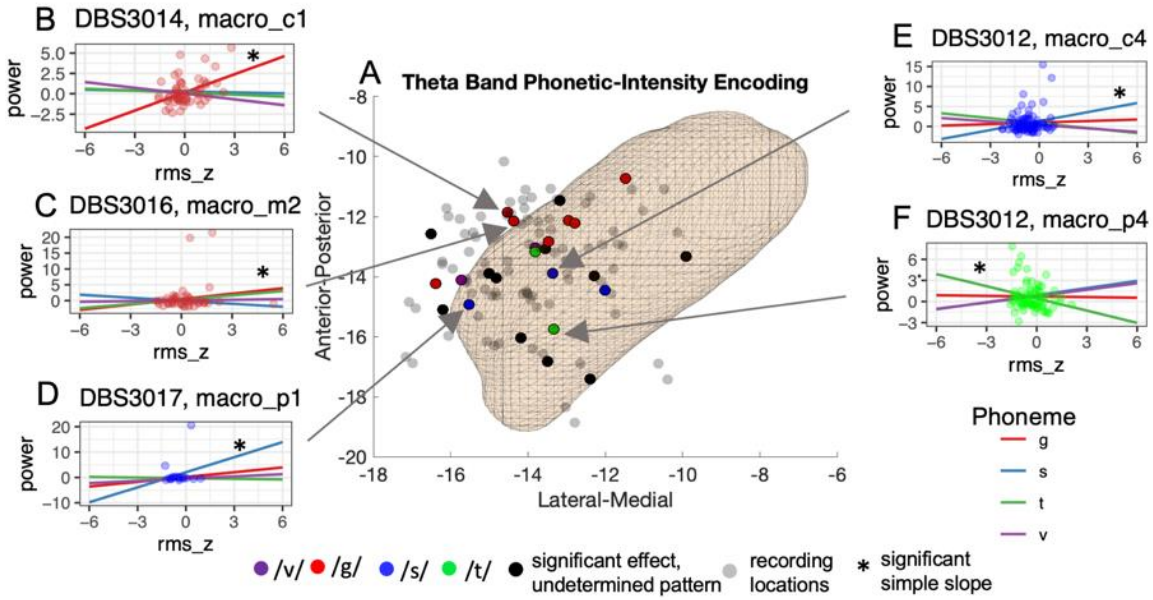


Figure 13. Phonetic-Intensity Encoding in the Subthalamic Nucleus

Phonetic-Intensity Encoding in the Subthalamic Nucleus (STN). A) Central plot depicts recording locations in the STN and site at which both phoneme and vocal intensity predicted neural power (theta band) via an interaction effect. B-F) Specific interaction effects for recording locations are depicted in the surrounding figures following simple slope analysis. Raw data points are included as scatter plots for each significant simple slope, color coded to match the slope of interest. Effects were found across participants; participant number is noted by title on surrounding figures.

Mixed-effects models predicting precentral gyrus power did not demonstrate a significant phoneme by electrode by intensity interaction effect; however, two-way interactions were significant for phoneme by electrode ($F_{(1115,\ 218510)} = 4.9309$, $p < .001$) and intensity by electrode ($F_{(372,\ 217840)} = 3.633$, $p < .001$) for gamma band power. No significant effects were found for theta band power in the precentral gyrus. These results indicate that specific recording locations are

involved in the component parts of *phonetic-intensity encoding*, although not overlapping in single locations as in the subthalamic nucleus.

6.3.4 Precentral Vocal Intensity Encoding Dependent on Articulator Regions

To determine if vocal intensity production in the precentral gyrus was influenced by articulator region, rather than specific recording sites or phoneme, recording sites were labeled by a specific articulator and then used to predict power, replacing the recording location term from the primary analysis. A mixed-effects model predicting neural power revealed significant two-way interaction effects for the lip by electrode ($F_{(357, 215933)} = 5.9$, $p < .001$), and larynx by electrode ($F_{(357, 215933)} = 3.2$, $p < .001$) terms. The tongue electrode variable was orthogonal to the lip condition; therefore, significant negative estimates for the lip condition were designated as tongue electrodes. Recording locations which significantly increased power for a specific articulator were then labeled as that articulator (Figure 14 A); results were verified by examining articulator label performance (Appendix B). A total of 31 recording locations were identified as having articulator specificity. These electrodes were then used to predict neural power with vocal intensity and phoneme. The three-way interaction effect between intensity, phoneme and articulator was not significant, consistent with the original electrode location analysis. However, a significant articulator by vocal intensity interaction was identified ($F_{(2, 22530)} = 5.8$, $p = .002$), indicating that as vocal intensity increased, articulator regions differentially responded (Figure 14 B). The larynx articulator regions decreased neural power as vocal intensity increased (simple slope estimate = -0.03, $p = .24$), while lip and tongue articulators increased power vocal intensity as vocal intensity increased (simple slope estimate lip = 0.02, $p = .29$; simple slope estimate tongue = 0.05, $p = .38$).

While these simple slopes were not independently significant, the interaction demonstrates a potential differential contribution of articulators to vocal intensity production.

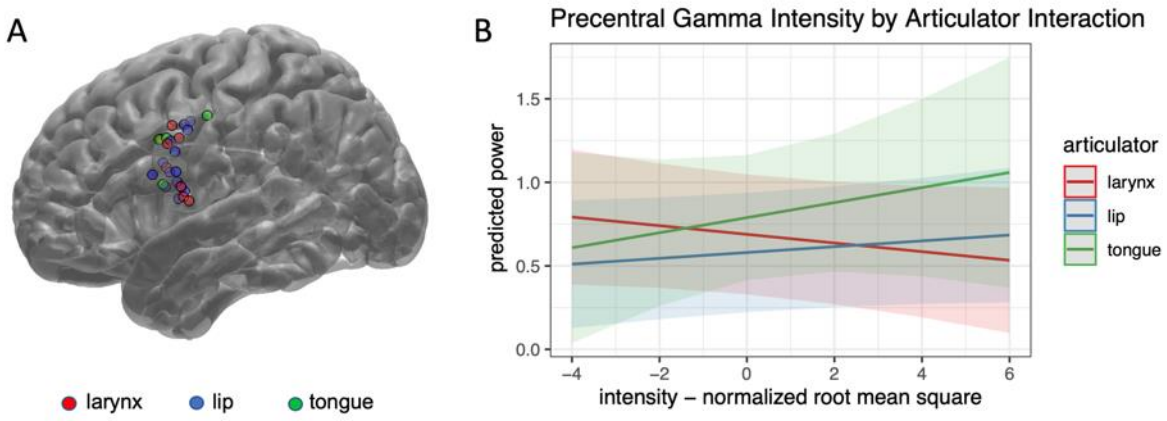


Figure 14 Articulator Regions Contribute to Intensity Control

A) Cortical locations which demonstrated a significant increase in gamma band related to a particular articulator. Articulator predictions were determined using a regression model, which predicted neural power from a dichotomous (0 or 1) variable for articulator usage. B) Gamma band power was predicted by vocal intensity and articulator region via a mixed-effects model. A significant articulator by intensity interaction was found, indicating differential response of articulator regions based on vocal intensity increases.

Disease severity, measured by Unified Parkinson's Disease Rating Scale did not predict neural power, indicating that findings are not attributable to the disease process.

6.4 Discussion

These data demonstrate the first neurological correlate of *phonetic-intensity encoding,* a theoretical process which is integral to clinical speech intervention (Dromey & Ramig, 1998; L. O. Ramig et al., 2008; L. O. Ramig, Sapir, Countryman, et al., 2001), but omitted from current theoretical accounts of speech production (Guenther, 2015; Levelt et al., 1999). Multiple neurological recording locations demonstrated overlapping representation of phoneme and vocal intensity production in the subthalamic nucleus, providing a neurological substrate for *phonetic-intensity encoding*. Prior electrophysiological (Dastolfo-Hromack et al., 2021; Lipski et al., 2018) and behavioral (Aldridge et al., 2016; Murdoch, 2001) research has postulated that the basal ganglia, and subthalamic nucleus specifically, are involved in speech functions. Data from the present work builds on this evidence by testing a specific speech motor control function and provides a topographical report of regions involved in *phonetic-intensity encoding* within the subthalamic nucleus and precentral gyrus. These data also contribute to a growing body of evidence demonstrating the instrumental role that neural oscillatory frequency bands may play in the control of speech production (Arnal & Giraud, 2012; Giraud & Poeppel, 2012; Hyafil, Fontolan, Kabdebon, Gutkin, & Giraud, 2015), specifically that theta and gamma band are critical neurological components to successful speech motor control. These results extend beyond interactions between phoneme and vocal intensity and tested the potential overlap between established articulator maps and vocal intensity control. Data suggest that vocal intensity control may also interact with regions that support speech articulator movements. These findings reveal that common speech functions may be neurologically represented in an overlapping manner and provide neurological evidence to more broadly understand the neural drivers of speech production.

Participants Improved Articulation with Increasing Intensity

Prior literature has consistently demonstrated that patients improve articulation with increasing vocal intensity (Dromey et al., 1995; Gates, 2020; Maniwa et al., 2009; Tjaden et al., 2013; Tjaden & Martel-Sauvageau, 2017) and the present results replicate these findings in the present cohort and experimental conditions. The data provided evidence from both temporal and spectral measurements that patients adjusted articulation. These findings of reduced frication time and prolonged vowel production (Figure 11) agree with effects found after Lee Silverman Voice Treatment (Dromey & Ramig, 1998; Dromey et al., 1995). Similarly, the spectral measures of increased spectral centroid and second formant ratio also agree with effects found after loud speech instruction (Tjaden & Martel-Sauvageau, 2017; Wenke, Cornwell, & Theodoros, 2010). The increase in the second formant ratio, specifically, is reflective of a more distinct speech articulation pattern (Sapir et al., 2010). The increase in spectral centroid likely indicates a more forward placement of the tongue during articulation, while the second formant ratio reflects greater anterior-posterior movements to create distinctions between forward /i/ and back /u/ vowel sounds (Kent & Read, 2002).

The effect sizes of these measures were small, whereas the intensity driven effects observed in normal populations (Dromey & Ramig, 1998) and effects after Lee Silverman Voice Treatment were moderate to large effect sizes (Wenke et al., 2010). Experimental conditions for these procedures were very different from the conditions of these prior studies; patients were lying supine and remained partially under the effect of anesthesia, although sufficiently awake to perform the task. The ideal bimodal distribution of vocal intensity values (for loud vs. typical speech) was not achieved, which may have influenced the size of these effects. Time constraints dictated by clinical care limited the amount of pre-experiment practice for vocal intensity

modulation, and the effects were likely influenced by anesthesia. Considering these conditions, the small, but significant changes observed in speech production are quite meaningful. Importantly, these effects confirm that this cohort produced the anticipated behavioral effects of *phonetic-intensity encoding*, a critical finding needed to assess neurological data.

Phonetic Intensity Encoding in the Subthalamic Nucleus

The basal-ganglia are integral to generalized movement control (Basinger & Joseph, 2021; Obeso, Rodriguez-Oroz, Rodriguez, Arbizu, & Gimenez-Amaya, 2002) and the presented data confirm that this role extends to speech production, specifically, *phonetic-intensity encoding*. Multiple locations in the subthalamic nucleus demonstrated vocal intensity-dependent increases in theta band power, but only for a single phoneme (Figure 13). This finding confirms the hypothesis that *phonetic-intensity encoding* occurs within the basal-ganglia cortical loop. It was surprising that evidence of *phonetic-intensity encoding* was not found in the precentral gyrus, given the previous findings of articulatory encoding (Bouchard & Chang, 2014a, 2014b; Brown et al., 2009). Presented results replicated the finding of phoneme-specific representation in the precentral gyrus (Bouchard & Chang, 2014b; D. Conant et al., 2014; Mugler, Goldrick, & Slutzky, 2014) as well as articulatory encoding, providing confidence that these analyses were adequately sensitive to identify relevant effects. It is possible that the precentral gyrus is responsible for final stage, detailed instruction of movement trajectory information, and may not be responsible for synthesis of multiple movement parameters. The finding that *phonetic-intensity encoding* was identified only identified in the STN and not precentral gyrus highlights the importance of the basal ganglia in speech production control.

Theories of basal ganglia function and primate literature support and inform the presence of *phonetic-intensity encoding* in the STN. Viral tracing studies in primates have found small, lateralized regions responsible for orofacial movement and primitive vocalization in the STN (Nambu et al., 1996). The present result from human participants indicates that vocalization regions remain in STN human anatomy but are more specialized in function and broadly distributed rather than present in only lateral regions (Figure 13). This topographical broadening and functional specification may indicate an evolutionary expansion of the function of the STN for speech production in humans.

Phonetic-intensity encoding underlies speech treatment for patients with Parkinson Disease and although this study was not a treatment study, it can partially inform the mechanisms that are utilized in treatment because it shares the same instruction, "speak loud." LSVT outcome studies have shown a shift in cortical activation from left to right following treatment (Liotti et al., 2003), including an increase in right precentral gyrus activation (Narayana et al., 2010). This was attributed to and increase in sensory integration areas caused by loud training. Data from the present study focused on left precentral gyrus and STN and did not explore right hemisphere function, which may potentially play a role in *phonetic-intensity encoding.* LSVT outcome studies have also found increased activation in other basal ganglia structure (Liotti et al., 2003), supporting a role for cortico-basal ganglia loops in vocal intensity control. This combined evidence suggests that phonetic effects of loud instruction on articulatory control (Dromey et al., 1995; Martel Sauvageau et al., 2015; L. O. Ramig, Fox, & Sapir, 2007; Sapir et al., 2010) are mediated by the basal ganglia, which may play a role in LSVT mechanisms. Therapeutic instructions which require a patient to 'speak loudly' likely activate the basal ganglia network. Finally, clinical speech outcomes after STN deep brain stimulation have been variable, with some studies explaining

changes in speech function on high voltage stimulation affecting motor tracts outside of the STN. However, the evidence of speech representations within the STN present the possibility that speech changes after STN-DBS may be caused by stimulation of speech centers within the STN itself. Future research is needed to determine examine these clinical outomes in more detail.

Vocal Intensity Encoding Overlaps with Articulator Features

The discovery of *phonetic-intensity encoding* illuminates the neural synthesis of phonetic and prosodic domains; however, it may not be the only method by which speech synthesis occurs. The presented results revealed another potential method of speech synthesis: cortical articulatory regions respond differentially to increases in vocal loudness. An interaction effect was found in the precentral gyrus, showing that as vocal intensity increased, gamma band activity in larynx regions reduced cortical power, while regions of the lip and tongue increased power (Figure 14). On initial inspection, it appears that this finding is counterintuitive because voicing, which is produced in the larynx, contributes to increases in sound pressure level (Baker, Ramig, Sapir, Luschei, & Smith, 2001). However, speech physiology research has shown that increases in vocal intensity are primarily mediated by a respiratory component, alveolar pressure, rather than substantial activation of intrinsic laryngeal musculature (Finnegan, Luschei, & Hoffman, 2000). Increases in alveolar pressure induce higher vocal intensity by causing a passive increase in the mucosal wave and medial-lateral displacement of the vocal folds, which may result in reduced neural activation of some laryngeal muscles. Furthermore, these data show that changes in articulation occurred with increasing intensity and it is possible that patients were producing clear speech as a strategy to increase vocal intensity, leading to increases in the articulators underlying the observed acoustic results (Figures 11 & 12). These articulatory-respiratory responses

demonstrate a emerging neurological pattern, that articulatory/linguistic functions and prosodic functions may be produced through overlapping representations in the brain.

Current theoretical accounts propose that neurological links between segmental and prosodic functions are accomplished through cortico-cortical connections (i.e. corpus callosum) (Klouda et al., 1988; Sammler et al., 2010); therefore, the discovery of regions with dual representation, both in *phonetic-prosodic encoding* and cortical articulatory-intensity encoding, provides a novel neural pathway for coordination between speech behaviors. Current models treat linguistic and prosodic parameters as separate systems (Guenther, 2015; Levelt et al., 1999), yet this compartmentalization may not adequately reflect neural processing. It is not surprising that neural substrates were identified connecting articulators with vocal intensity, because behavioral accounts have described a reciprocal pattern between articulators and laryngeal movements (i.e. laryngeal articulatory coupling) (Dromey, 2010). These findings of dual functioning also align with the architecture of basal ganglia-cortical loop. Primate literature has demonstrated that movement effectors are represented in both the motor cortex and basal ganglia, and these somatotopic representations are anatomically connected between brain regions (Alexander et al., 1986; M. DeLong & Wichmann, 2010). These results suggest that this type of functional organization may also apply to speech production. Thorough analysis of the STN articulator regions was not possible due to a lack of diversity in the articulators identified (primarily lip labeling). This would be a valuable area for future research. Taken together, these findings of overlapping representation for articulation and intensity as well as *phonetic-intensity encoding* indicate the need to broadly consider how simultaneous linguistic and prosodic functions are seamlessly accomplished in speech production.

Theta and Gamma are Critical Bands for Speech Production.

Current neurological speech production theory identifies gamma and theta band as critical components of speech production (Giraud & Poeppel, 2012) and the present results add to this growing evidence base. Giraud and colleagues found that mouth movements correlate to gamma band power in the premotor area (Giraud et al., 2007), which agrees with the present finding of articulator encoding in precentral gamma band. These data identify not simply mouth movements but also overlapping articulatory-intensity functions in premotor gamma band. *Phonetic-intensity encoding* also extends the proposed theta band modulated cortical regions to include subcortical theta band activity. Subcortical (STN) theta band in these data showed overlapping representation of speech features; however, the current theoretical hypothesis is that gamma and theta band are responsible for coordination of the complex timescale of movements (Giraud & Poeppel, 2012). Acoustic temporal changes were produced in vowels and fricative with increasing intensity (Figure 11), and it is possible that gamma and theta band mediated these effects.

The pattern of activity, namely gamma band in cortex and theta band power in the STN, may not be coincidental. Canolty and colleagues found that gamma band power phase-locked to theta oscillations (Canolty et al., 2006). Similar types of coherence relationships have been found between theta and gamma band for speech production, including phase-amplitude coupling (Lizarazu et al., 2019) and theta-modulated gamma band (Doesburg et al., 2012). It is possible that these two frequency bands work together within the basal ganglia cortical loop to modulate numerous aspects of speech production, including *phonetic-intensity encoding.* Altogether, these findings support the concept that theta and gamma band power are critical modulators of speech production.

Conclusions

Synthesis between speech domains of articulation and prosody (intensity) occur within the basal ganglia-cortical loop. The most salient is *phonetic-intensity encoding* in the STN, which is relevant for speech treatment in Parkinson Disease. Preliminary evidence for a secondary synthesis mechanism is in left precentral gyrus, with articulator encoding modulating intensity encoding. Finally, these results highlight the importance of gamma and theta band in the neural control of speech production, specifically the fusion of linguistic and prosodic processing.

Limitations

These effects were found in patients with Parkinson Disease, and although the analysis of the UPDRS score concluded that disease severity did not predict the outcome measures, influence from the disease process is still possible. Articulatory encoding was completed on a limited set of phonemes, which limited the number of identifiable articulators.

7.0 Manuscript #2

As stated at the beginning of the manuscript #1, this section details the results from the question from the second aim which addresses the differential contributions of the precentral gyrus and subthalamic nucleus to vocal intensity encoding. Please note that portions of the methods section are identical to manuscript #1.

WORKING TITLE: Vocal intensity encoding in the basal-ganglia cortical loop.

7.1 Introduction

Speech production relies on varied and nuanced vocal intensity to produce prosody which communicates the speakers emotional or linguistic intention (Iredale, Rushby, McDonald, Dimoska-Di Marco, & Swift, 2013; Pichon & Kell, 2013) and this process requires precise neurological control. Studies of the neural origin of prosody primarily focus on cortical control and omit subcortical structures (Baum & Pell, 1999), like the basal ganglia, and how subcortical structures interact with cortical control. This omission is, in part, due to the difficulty in obtaining data from subcortical structures. Behavioral observations of speech characteristics in patients with Parkinson's disease (Tjaden, 2008) and Huntington's chorea (Ludlow et al., 1987) have led researchers to suspect that the basal ganglia are involved in prosodic production, prominently vocal intensity, indicating the need for more subcortical prosody research. Furthermore, the motor cortex and basal ganglia are considered critical regions for general motor control and are

functionally and anatomically connected in what is termed the 'basal ganglia-cortical loop'(Alexander et al., 1990; Alexander et al., 1986; M. DeLong & Wichmann, 2010). These interconnected structures offer a potential substrate for prosodic control. Understanding the role of the basal ganglia-cortical loop in the control of one component of prosody, vocal intensity, would contribute to neurological theories of speech motor control. The goals of this study were to 1) determine if any regions within the precentral gyrus or subthalamic nucleus were involved in vocal intensity control and 2) understand the differential contributions of these two structures to vocal intensity control.

Motor control studies in both primates and humans have identified the basal-ganglia-cortical loop as a critical modulator of limb movements. Modulations in movement force, speed and amplitude are correlated to electrophysiologic markers of gamma band (Anzak et al., 2013; Tan, Pogosyan, Anzak, Ashkan, et al., 2013) and theta band (Tan, Pogosyan, Anzak, Ashkan, et al., 2013). These types of movement alterations, although found in non-speech motor tasks, are highly relevant to vocal intensity production, which requires modulation of the amplitude and speed of respiratory and laryngeal movements (Isshiki, 1964). Recently, the acoustic amplitude of movement (articulatory gain) was correlated to alpha and theta band in a section of the basal ganglia, the subthalamic nucleus, and the precentral gyrus (Dastolfo-Hromack et al., 2021). These findings strongly suggest that vocal intensity production may be controlled in these structures, and that correlates to vocal intensity may be found in oscillatory frequency bands. Although these studies have contributed greatly to the understanding of motor control, most of this evidence comes from limb movements, which are theoretically different from speech, or arise from observational data. The present study will utilize a causal design to investigate vocal intensity representation within the basal ganglia-cortical loop.

Theories on the neural control of prosody emphasize cortical control (Baum & Pell, 1999), with prominent theories addressing cortical lateralization issues (Klouda et al., 1988; Mayer et al., 2002; Van Lancker, 1980) while ignoring the influence of subcortical structures. Some researchers argue that subcortical structures are largely responsible for prosodic production (Cancelliere & Kertesz, 1990). The present study will inform the subcortical hypothesis, as well as improve upon other limitations in prior research. Extant research either focuses on prosody as a generalized linguistic concept (Aziz-Zadeh et al., 2010; Baum & Pell, 1999; Caekebeke et al., 1991; Golfinopoulos et al., 2015; Mayer et al., 2002; Paul et al., 2003; Pichon & Kell, 2013; Riecker et al., 2002; Van Lancker Sidtis et al., 2006), or on the linguistic subcomponents of emotional vs. linguistic prosody (Pichon & Kell, 2013). However, fewer studies have investigated the neural basis of acoustic properties of speech, pitch, intensity and duration (Baum & Pell, 1999).

Clinical studies which measured acoustic properties of prosody support the hypothesis that vocal intensity is controlled by both cortical and subcortical structures. Diffuse changes in cortical activity (Narayana et al., 2010; Rektorova et al., 2007) and increased activation in the basal ganglia (i.e., caudate nucleus)(Liotti et al., 2003) were found while studying outcomes of a prominent speech therapy that modulates vocal intensity, Lee Silverman Voice Therapy (LSVT®). The basal ganglia are also implicated in vocal intensity because changes in vocal intensity (both increases and decreases) occur following deep brain stimulation (Aldridge et al., 2016; Behroozmand et al., 2019). These clinical studies offer some insight into vocal intensity production, but do not isolate vocal intensity from other clinical processes. Therapeutic outcomes may capture numerous changes that occur as part of treatment and may not be effects isolated to vocal intensity modulation.

These investigations of prosody and vocal intensity do not describe direct neural processes, but rather infer these processes from indirect measures (i.e. BOLD signal). Data from direct neural processes, such as changes in neural population dynamics, would not only inform hypotheses regard prosody control, but also provide biomarkers which may be used in future interventions, such as deep brain stimulation. Nascent studies of deep brain stimulation describe differences in speech outcomes based on stimulation frequency within the STN (Aldridge et al., 2016). Intelligibility and speech naturalness has been found to decrease with high frequency stimulation, compared to low frequency stimulation (Grover et al., 2019; Tanaka et al., 2020). Greater understanding of the roles of high frequency (i.e. Gamma band 80-150 Hz) and low frequency (i.e. Theta band 4-8 Hz) oscillations in neural populations may produce hypotheses for improving speech outcomes after deep brain stimulation.

To further understand the neural basis of vocal intensity production, a protocol was developed to intraoperatively measure electrophysiological signals from the precentral gyrus and subthalamic nucleus during awake deep brain stimulation implantation. Patients were asked to modulate vocal intensity while speaking non-speech stimuli which were balanced for phonemes. Theta and gamma band power were extracted because of their importance in speech motor control (Giraud & Poeppel, 2012; Hyafil et al., 2015; Sengupta & Nasir, 2016) and prior evidence of amplitude modulation in these structures (Dastolfo-Hromack et al., 2021). These data were used to predict changes in vocal intensity, while controlling for disease severity, to test the hypothesis that vocal intensity encoding occurs in the basal-ganglia cortical loop. Results indicate that both the subthalamic nucleus and precentral gyrus are involved in vocal intensity control; however, these regions may contribute differently to the vocal intensity production.

7.2 Methods

7.2.1 Participants.

Data were collected from a total of 29 patients presenting to a neurosurgical clinic for deep brain stimulation surgery. Patients were recruited during clinical visits during which they were undergoing evaluation for deep brain stimulation implantation surgery. Inclusion criteria were 1) a diagnosis of idiopathic Parkinson's disease, 2) recommendation for surgical treatment by a multidisciplinary movement disorder surgery team, which requires thorough neuropsychological evaluation and evaluation of PD symptoms on and off medication, 3) neurosurgical target of the STN, and 4) favorable anatomy free of venous obstructions that would impede electrode contact with the cortical surface. Patients were excluded if they were not candidates for deep brain stimulation based on the above criteria or were unable to perform the task. To ensure informed consent occurred without coercion, study participation meetings and informed consent procedures were held with a research coordinator who was not involved in clinical processes. Patients were assured that the decision to participate would not affect their clinical care in any manner. This research study was approved by the University of Pittsburgh Institutional Review Board (PRO13110420) and all participants completed a written informed consent.

7.2.2 Stimuli

Three syllable 'triplets' (e.g. 'see too gah') were constructed from different combinations of 4 consonants (/g, t, s, v/) and 3 vowels (/i, a, u/). These phonemes were chosen because they consist of articulatory features (involvement of the lips, tongue tip, low tongue/jaw, front tongue, and

larynx) and phonetic state spaces (high, low, front, and back vowels; alveolar, labial, and coronal consonants) which have previously been described in the human motor cortex. Phonemes were presented pseudo-randomly, such that the initial phoneme of the triplets was balanced, and each consonant was presented in the initial position 30 times within a session. Participants were also auditorily prompted (by acoustic example) to increase intensity on half of the trials. The intensity increase was balanced within phoneme condition, so that in a single phoneme condition, 15 were normal intensity and 15 were increased intensity (increase of 20 dB SPL). Actual intensity level of the acoustic stimulus presentation varied based on patient, and audio was adjusted to achieve comfortable hearing level for each participant. The magnitude between loud and soft productions was constant.

Table 5. Stimuli Characteristics

Consonant	Alveolar	Labiodental	Velar	Fricative	Voicing
/v/	0	1	0	1	1
/t/	1	0	0	0	0
/s/	1	0	0	1	0
/g/	0	0	1	0	1

7.2.3 Pre-operative Training

While in pre-operative care, patients completed a pure-tone hearing screening at 500Hz, 1kHz, 2kHz, and 4kHz at an intensity level of 25- and 40-dB HL using a calibrated tone generator. Initial hearing screenings were completed by the PI (licensed speech-language pathologist) and then a trained research assistant. Patients were instructed that some of the sounds would be soft, and some are loud, and to use a strong, loud voice when prompted by the stimuli. Patients completed practice using the stimuli and were coached as necessary.

7.2.4 Surgical Procedure and Electrode Placement

Prior to surgery, patients were off medication for 12 hours. Routine clinical procedure dictates that, while patients are anesthetized, a surgical frame is placed around the head for stabilization and burr holes are drilled into the skull at the point of safest entry. After waking the patient from anesthesia, a micro-electrode recording (MER) array, consisting of three electrodes in a fixed triangular orientation are inserted into the brain via the burr hole; three macroelectrode contacts are located superior to the macroelectrodes. LFP recordings were gathered from these macroelectrodes. Trajectory of the MER is guided by pre-operative imaging, and continuous electrophysiologic monitoring by a neurophysiologist. Single cell recordings are used to determine location of the electrode at the top or bottom of the STN. Recording depths are extrapolated using the changes in depth during placement measured on the MER array, combined with depths associated with STN top and bottom (determined by presence of neuron firing patterns associated with STN). Cortical high density 64 contact electrode strips were also placed through the burr onto the cortical surface, targeted toward the sensorimotor cortex. Trajectory of the cortical electrode strip was planned pre-operatively using fMRI images, and a corresponding marker was placed on the patient's scalp during surgery to guide placement. After electrodes were in place, a fluoroscopy image was taken for post-operative localization. Patients were instructed to repeat acoustic stimuli, which were presented through earbuds. Recordings were taken at 2-4 depths within the STN, corresponding to 2-4 sessions depending on patient fatigue and surgical concerns. One session consisted of 120 nonsense syllable 'triplets'.

7.2.5 Intra-operative Recording

Local field potentials were recorded from the deep brain stimulation lead after placement, and this was completed using the Grapevine neural interface processor (Ripple LLC, Salt Lake City, UT, USA). Cortical electrode strips were microphonic-free to reduce movement artifact. Stimuli of triplets were delivered acoustically to the patient through ear buds (ER38-14F foam tips; Etymotic Research, Inc.) after the signal was amplified (PreSonus, AudioBox iTwo). Audio recordings were gathered using a high fidelity recorder (H6 Handy Recorder, Zoom Inc., Japan) at a sampling rate of 96kHz, using an omnidirectional microphone in order to reduce distortion related to vocal intensity (MSH-6 Stereo Mic, Zoom Inc., Japan), at 7 cm, 45^{o} angle from the oral angle.

7.2.6 Pre-processing

All electrophysiological data were processed in MATLAB with custom scripts using the Fieldtrip toolbox (Oostenveld et al., 2011). Data were resampled to a sampling rate of 1 kHz, and cardioballistic effects and line noise were removed using a 1Hz high pass filter and 58-62 Hz notch filter. A high pass Butterworth IIR filter ('but') was applied to raw data with a zero-phase forward and reverse filter ("two pass"), with a filter order of 5. High frequency movement artifacts were identified algorithmically by first calculating the log-normal distribution of the signal amplitude, then determining an appropriate cut off criteria for each participant. The artifact cut off was between 2-3 standard deviations away from the median of the distribution and was determined by examining the distribution and placing the cut off at the point where the tail of the distribution began. Artifact detection was individualized per participant. In addition, a speech production vibration related artifact was identified, and any session determined to be affected by the vibration

related artifact was removed from analysis based on two published criteria (Bush et al., 2021). For cortical electrodes, LFP signals were re-referenced offline to a common average reference, applied over blocks of electrodes. Because only three contacts existed for STN recording sessions, a bipolar referencing scheme was applied to the STN contacts. Data were aligned to the onset of each consonant for epoching. Wavelet transformation was performed using a custom MATLAB function. Wavelet parameters were a width of 7 cycles, evaluated at theta (4-8 Hz) and high gamma frequencies (70-150 Hz)(Lachaux et al., 2007; Nir et al., 2007). Power values were extracted from the wavelet transformation using a custom function (BML lab toolbox)(Bush, 2021) and baseline normalized on a trial-by-trial basis. A baseline period was designated as one second prior to the onset of the stimulus. Epoching and baseline procedures were repeated for both trial level data and syllable level data to support the describe statistical analysis. Power from each trial or syllable segment was normalized to the appropriate baseline for the trial from which it was obtained. Boxplots of normalized power values was examined for outliers. Across participants, most outliers appeared above 2.5 standard deviations above the mean (amplitude of 30). A random sample of these outliers was visually examined, and all represented residual movement artifact; therefore, normalized power values above 30 were excluded from analysis.

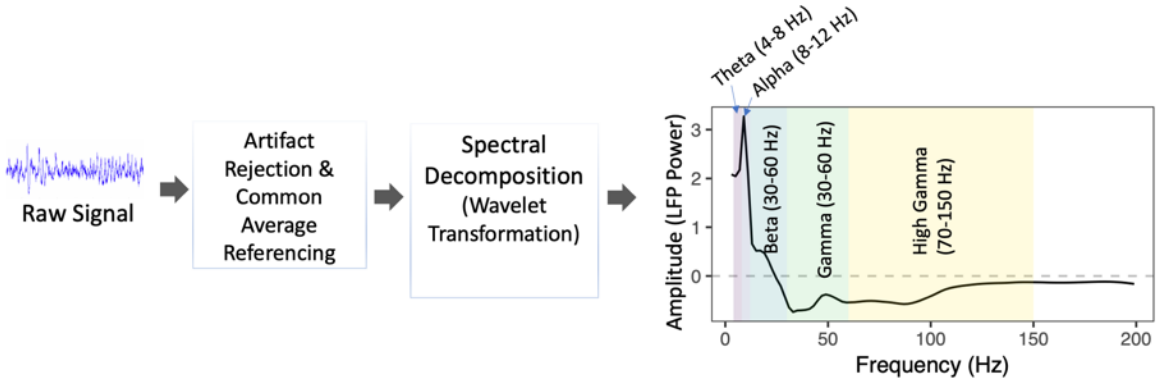


Figure 15. Data Processing Flowchart.

Overview of the processing required to extract the final power values from the raw electrophysiologic signal, which was used as the primary neural measurement.

7.2.7 Electrode Localizations.

Cortical electrode recording locations were determined from reconstructions using intraoperative fluoroscopic images, co-registered preoperative and post-operative computed tomography (CT) images and pre-operative magnetic resonance imaging (MRI) scans, procedures are described in Randazzo et al. (2016). Localization of electrodes into patient native space were conducted using the Freesurfer suite (Dale et al., 1999) (http://surfer.nmr.mgh.harvard.edu/) and a custom-made graphical user interface in MATLAB. These localizations were registered into the MNI common space template using Brainstorm (Tadel et al., 2019) (https://neuroimage.usc.edu/brainstorm/). Reconstruction of the subthalamic nucleus recording trajectory was conducted using the Lead DBS toolbox in MATLAB, (Horn & Kuhn, 2015) based on identification of lead-induced artifact in the post-operative CT scans.

7.2.8 Coding of Speech Data

Phonetic transcription of each participant's produced speech and the markings of speech onset were performed by a trained team of speech pathology students. Using a custom Matlab GUI (github.com/Brain-Modulation-Lab/SpeechCodingApp) (example in Appendix A), onset and offset times were marked based on acoustic evidence of key speech features. Voiced phonemes (i.e. /v/) and vowels were marked at the first visible glottal pulse, indicating the onset of vocal fold vibration. Phonemes that do not rely primarily on voicing were marked based on the characteristic noise features for that phoneme. For example, the rapid increase of high frequency energy denoted the onset of plosive consonants (i.e. /t/ and /g/). A broad band spectrogram was utilized to maximize visualization of key speech features. All coders completed a course in speech science to ensure knowledge of important speech features and were trained in coding procedures by a speech language pathologist. Additional information on coding procedures is available in Appendix A.

7.2.9 Speech Performance Measurement

Vocal intensity level was measured by calculating the root mean square (RMS) per utterance and per syllable. One syllable consisted of a consonant-vowel pair and the entire utterance consisted of three syllables from a single trial. Normalized (mean z-scored) RMS was calculated for intensity within participants. For equation 2, n = number of samples in the syllable, x = raw amplitude value at a single sample, μ = mean RMS per participant, σ = standard deviation of RMS values per participant.

Equation 3

$$Equation:\ RMS\_Z = \frac{\sqrt{\sum \frac{x^2}{n}} - \mu}{\sigma}$$

7.2.10 Parkinson Disease Severity Measurement

To control for disease severity, the Movement Disorder Society-Unified Parkinson Disease Rating Scale-III (UPDRS) and time since diagnosis onset were gathered. The UPDRS was administered as part of the pre-clinical measurements for deep brain stimulation surgery, during which the electrophysiology measurements were gathered. The UPDRS was administered by a neurosurgery physician's assistant with specialized training in movement disorders and significant clinical experience with Parkinson's disease patients.

7.2.11 Statistical Analysis

Vocal Intensity Modulation

To determine if participants modulated vocal intensity, a t-test was calculated comparing the productions which were instructed to be soft compared to those instructed to be loud. Calculations were completed on the normalized vocal-intensity data. In addition to calculation, the vocal intensity distribution was examined on a per-participant basis to evaluate the extent of vocal intensity modulation.

Neural Correlates to Vocal Intensity Production

To broadly evaluate the role of the precentral gyrus, subthalamic nucleus and oscillation frequency band in vocal intensity control, a mixed-effects model was fit, predicting trial level vocal intensity from the fixed effects of STN power, precentral gyrus power, neural band, and measures of disease severity at two time-points. Two models were fit: 1) before speech onset during the reaction time interval to broadly assess predictions during motor preparation and 2) during speech production to

assess motor execution. The reaction time interval was defined as the average time between the stimulus onset and onset of speech production (latency), while the speech production interval was selected as the onset of speech production to the average duration of speech production. These two time frames were selected in order to examine the potential differences in relationship during motor preparation and motor execution. Power data were first averaged across participant and session, collapsing across electrode recording location in the precentral gyrus. Data were then averaged across time points during reaction time and execution intervals based on the average speech latency (0.72 seconds) and average speech duration (1.2 seconds). Participant was included as a random effect in the mixed-effects models. A second linear model was conducted evaluating disease severity predictors alone. Disease severity measurements are singular measurements per individual, causing the values to be replicated to fit the number of neural measurements; a second analysis of disease predictors along guards against alpha inflation. For the equation below: Region = subthalamic nucleus or precentral gyrus, Band = oscillatory band factor (gamma and theta) UPDRS = unified Parkinson disease rating scale, TSD = time since diagnosis.

$$
\begin{aligned}
&Participant\ Vocal\ Intensity\ (dB\ SPL) \\
&\quad = \gamma_1(Power) + \gamma_2(Region) \\
&\quad + \gamma_3(Band) + \gamma_4(Power * Region) + \gamma_5(Band * Power) \\
&\quad + \gamma_6(Band * Region) + \gamma_7(Power * Region * Band) + \gamma_8(UPDRS) \\
&\quad + \gamma_4(TSD)
\end{aligned}
$$

Location-Dependent Modulation of Vocal Intensity

To further understand effects found in the averaged data, a second within-subjects mixed-effects model was fit to evaluate topography of the vocal intensity effects as well as account for the variables of phoneme production and syllable. Neural power was predicted by syllable level vocal intensity, phoneme, and recording location. A maximal model was fit, including trial and subject as random effects. Four total models were conducted, to assess the topography in both gamma and theta bands, in the precentral gyrus and STN.

$$
\begin{aligned}
Neural\ Power \\
&= \gamma_1 Vocal\ Intensity + \gamma_2\ Phoneme + \gamma_3 Location \\
&+ \gamma_4 Vocal\ Intensity\ x\ Phoneme + \gamma_5 Vocal\ Intensity\ x\ Location \\
&+ \gamma_6 Phoneme\ x\ Location + \gamma_7 Vocal\ Intensity\ x\ Location\ x\ Phoneme
\end{aligned}
$$

Following model completion, a type III analysis of variance was conducted to assess overall effect significance for each main effect. The primary effects of interest are γ_1 and γ_5, predictions of vocal intensity and the interaction between vocal intensity and electrode location. Significance in either of these terms indicates an effect of vocal intensity production within the region and neural band tested, with the interaction effect indicates that intensity is represented differently across recording locations. A Bonferroni corrected alpha of 0.0125 was utilized to account for the same hypothesis being tested in four different models.

Differential Time-Course of Vocal Intensity Control

After determining the electrode locations which predicted vocal intensity, a qualitative analysis was completed to assess the time-course of power modulation. The normalized power data were

examined in the electrodes which predicted vocal intensity in the syllable-level analysis. For these electrodes, trials were split into two group, high and low vocal intensity, determined by the median vocal intensity value for the session. The averaged neural data were then plotted across all timepoints in the task and separated by vocal intensity level for visual analysis.

7.3 Results

7.3.1 Participants and Behavioral Performance

Intraoperative and acoustic recordings were gathered from 29 patients with Parkinson's disease undergoing STN deep brain stimulation; however, 9 were eliminated due to lack of recordings in the precentral gyrus or based on artifact rejection criteria leaving a total of 20 patients. Patients were an average age of 65.5 years, with the majority (72%) reporting male gender and right-hand dominance (79%), which is typical of the greater population of patients with Parkinson's disease. Severity of disease was mild-moderate, with an average UPDRS score of 33.7 and average duration of disease reported was 6.5 years. Please see Tables 6 and 7 for patient demographic information. Vocal intensity data were examined for contamination by background noise and non-speech patient noises, and only clean data were included for analysis; data elimination led to less than 1% data exclusion. Participants produced an average vocal intensity of 0.11 RMS with an average range of 0.26 RMS (Table 7). Vocal intensity production was significantly different between instructed soft trials and instructed loud trials (Normalized Mean Difference = 0.033, $p = .0075$), indicating that patients were completing the task as instructed. However, patients produced a range of vocal intensities (Figure 16) with a normal distribution,

rather than a bimodal distribution. Because vocal intensity variability did not naturally occur in two distinct groups, the data were not treated as two separate productions (loud vs. soft), but rather the entire range of vocal intensities was used in analysis to predict neural power in subsequent analyses.

Table 6. Participant Demographics

Participant	Gender	Age	Handedness	Hoehn & Yahr Stage	Duration of disease, years	UPDRS Speech Score	UPDRS Total Score
3001	male	69	right	2	6.5	0	25
3002	male	60	right	NA	5.6	1	29
3003	male	51	left	NA	6.8	1	17
3004	male	61	right	2	10.8	2	61
3005	male	59	right	2	4.8	1	46
3006	female	68	right	2	4.9	1	9
3008	male	70	right	NA	5	NA	NA
3010	male	70	right	NA	1.1	1	34
3011	male	66	right	NA	8.2	1	30.5
3012	male	70	right	NA	20.3	2	43
3014	male	71	unknown	NA	15.4	0	30
3015	male	75	right	NA	6.4	1	42
3016	female	66	right	NA	8.7	2	45
3017	male	76	right	3	0.8	1	34
3018	female	62	right	NA	1.9	1	21
3019	male	71	right	NA	6.9	NA	40
3020	male	64	right	NA	9.3	0	19
3021	female	50	right	NA	4.8	0	6
3022	male	67	right	NA	5.1	1	38
3023	male	61	right	NA	3.7	2	30
3024	female	63	right	NA	5.3	1	37
3025	female	74	unknown	NA	4.5	1	22
3026	male	69	right	2	10.6	1	50

3027	male	57	left	2	3.6	0	33
3028	female	64	left	NA	4.1	2	39
3029	male	78	unknown	NA		NA	NA
3030	female	62	right	NA	8.7	2	43
3031	male	70	right	NA	4.7	2	39
3032	male	56	right	NA	5.8	1	47

Table 7. Patient Performance and Recording Information

Participant	Number of Precentral Electrodes	Number of STN Contacts	Mean Vocal Intensity (raw root mean square)	Speech Latency (seconds)	Three Syllable Duration (seconds)	Number of Sessions
3001	NA	9	0.09 (0.03)	0.63 (0.22)	1.25 (0.21)	4
3002	10	6	0.04 (0.01)	0.44 (0.18)	0.96 (0.10)	3
3003	16	5	0.09 (0.04)	0.90 (0.24)	1.38 (0.14)	4
3004	NA	7	0.18 (0.05)	0.69 (0.26)	1.36 (0.20)	3
3005	NA	NA	0.06 (0.03)	0.58 (0.16)	1.06 (0.17)	3
3006	12	NA	0.21 (0.07)	0.85 (0.23)	1.64 (0.14)	4
3008	7	6	0.08 (0.08)	0.70 (0.21)	1.51 (0.12)	3
3010	35	9	0.02 (0.01)	0.76 (0.33)	1.16 (0.11)	4
3011	16	5	0.07 (0.03)	0.55 (0.17)	0.88 (0.10)	4
3012	17	12	0.15 (0.03)	0.79 (0.18)	1.17 (0.09)	5
3014	6	3	0.15 (0.05)	0.57 (0.33)	1.35 (0.27)	4
3015	7	2	0.05 (0.02)	0.83 (0.23)	1.60 (0.11)	4
3016	10	6	0.13 (0.04)	0.77 (0.57)	1.12 (0.16)	2
3017	24	4	0.26 (0.12)	0.70 (0.67)	1.56 (0.28)	3
3018	17	3	0.15 (0.09)	0.89 (0.20)	1.28 (0.08)	3
3019	15	4	0.05 (0.02)	0.82 (0.25)	1.19 (0.17)	3
3020	49	6	0.01 (0.01)	0.86 (0.35)	1.29 (0.22)	3
3021	22	NA	0.07 (0.02)	0.82 (0.15)	1.29 (0.11)	3
3022	14	6	0.15 (0.09)	0.69 (0.21)	0.98 (0.10)	3
3023	52	NA	0.15 (0.04)	0.62 (0.27)	1.19 (0.25)	3
3024	NA	6	0.14 (0.05)	0.64 (0.11)	1.13 (0.10)	3
3025	NA	NA	0.11 (0.04)	0.66 (0.28)	1.27 (0.29)	3
3026	NA	NA	0.10 (0.03)	1.15 (0.45)	1.23 (0.13)	2
3027	NA	NA	0.06 (0.03)	0.72 (0.32)	1.32 (0.25)	3
3028	NA	NA	0.08 (0.03)	0.66 (0.13)	1.43 (0.09)	4
3029	NA	NA	0.16 (0.06)	0.63 (0.17)	1.38 (0.14)	2
3030	23	3	0.07 (0.02)	0.72 (0.15)	2.12 (0.58)	2
3031	NA	NA	0.18 (0.14)	0.56 (0.17)	1.3 (0.12)	2
3032	16	9	0.14 (0.06)	1.15 (0.46)	1.49 (0.12)	3

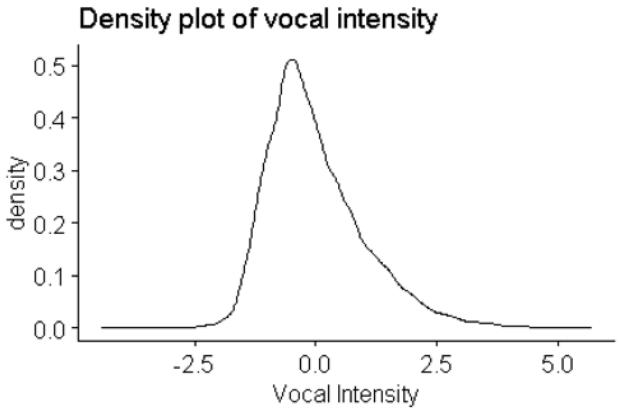


Figure 16 Density Plot of Vocal Intensity Productions

The density plot above reports the density of vocal intensity productions, measured in root mean square, for the respective distribution percentile. Data demonstrated a normal distribution, not bimodal; therefore, vocal intensity data were analyzed in a continuous fashion and not categorical.

7.3.2 Precentral Gyrus and Subthalamic Nucleus Predict Vocal Intensity

Mixed-effects models were fit predicting vocal intensity from averaged power in the precentral gyrus and STN, along with oscillation band and disease factors. Models were conducted during the reaction time interval and during speech production. Vocal intensity was significantly predicted by averaged neural power during the reaction interval ($F_{(1, 200)} = 5.38$, $p = .021$; $\beta = -0.315$, $p = .0391$); though that effect did not differ significantly between regions of frequency bands. During movement execution, a significant power by band interaction effect ($F_{(1, 200)} = 6.76$, $p < .01$) indicated that gamma band power significantly predicted vocal intensity in a negative direction across both brain regions ($\beta = -0.2$, $p = .02$)(Figure 17). Disease factors, UPDRS and time since diagnosis, did not demonstrate significant effects in either the mixed-effects model or separate linear model analysis.

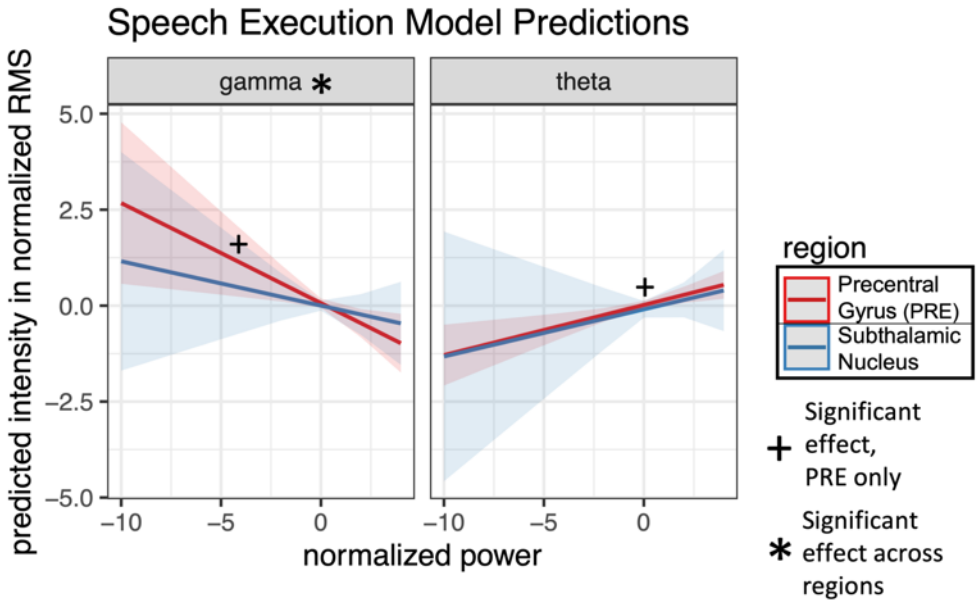


Figure 17 Vocal Intensity Production Predicted by Regions and Bands During Speech.

Mixed-effects models were fit, predicting vocal intensity production from neural power, band and neurological region (precentral gyrus or subthalamic nucleus). Data shown here were averaged across the speech production interval.

7.3.3 Differential Encoding of Vocal Intensity Across the STN and Precentral Gyrus

To assess vocal intensity effect topography and control for speech sound variability, four mixed-effects models were conducted predicting neural power from vocal intensity, in the precentral gyrus and STN with both theta and gamma band power. Gamma band power in the precentral gyrus was significantly predicted by both vocal intensity and electrode location via an interaction effect ($F_{(1115, 218510)} = 4.93$, $p < .001$). Theta band power in the STN was significantly predicted by the interaction effect between vocal intensity and electrode location ($F_{(110, 27958)} = 2.58$, $p < .001$), as well as the three-way interaction effect between vocal intensity, electrode, and phoneme ($F_{(330,}$

$_{27958)}$ = 1.39, $p < .001$). No significant effects were found in the STN gamma band or precentral gyrus theta band. Follow-up analyses indicated that four electrodes in the precentral gyrus positively predicted gamma band power from location and vocal intensity, while 35 electrodes negatively predicted gamma band power. For the STN, 39 electrodes positively predicted theta band power from location and vocal intensity, while one electrode negatively predicted theta band power (Figure 18).

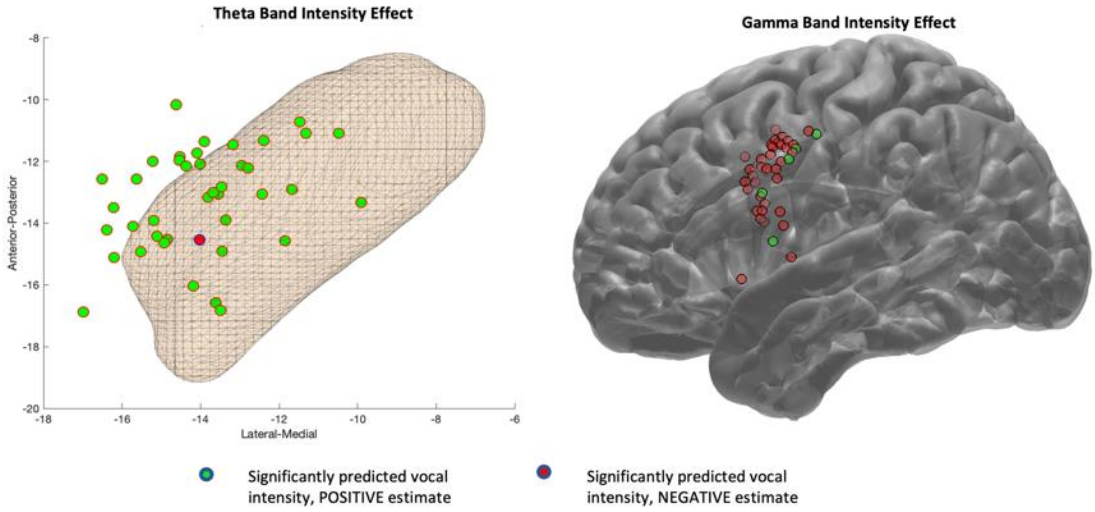


Figure 18. Topographical Differences in Vocal Intensity Encoding Across Regions

Phonetic-Intensity Encoding in the Subthalamic Nucleus (STN). Central plot depicts recording locations in the STN and site at which both phoneme and vocal intensity predicted neural power (theta band) via an interaction effect. Specific interaction effects for recording locations are depicted in the surrounding figures following simple slope analysis. Raw data points are included as scatter plots for each significant simple slope, color coded to match the slope of interest. Effects were found across participants; participant number is noted by title on surrounding figures.

7.3.4 Intensity Modulating Regions Express Differential Time-Course of Activity

To qualitatively assess the time-course of activity, neural power in electrodes which positively predicted vocal intensity was split between loudly produced trials and softly produced trials and visualized (Figure 19). Precentral gyrus electrodes which encoded for vocal intensity demonstrated a sustained difference throughout the entire task, with loud productions clearly exhibiting higher neural power compared to soft productions. Electrodes in the STN which positively predicted vocal intensity exhibited a variable time course of neural activity. Soft intensity productions demonstrated higher theta power both before and after speech production, but during speech production loud productions exhibited higher theta band power compared to soft productions. The time surrounding 0.4 seconds after speech onset displayed no difference between intensity levels.

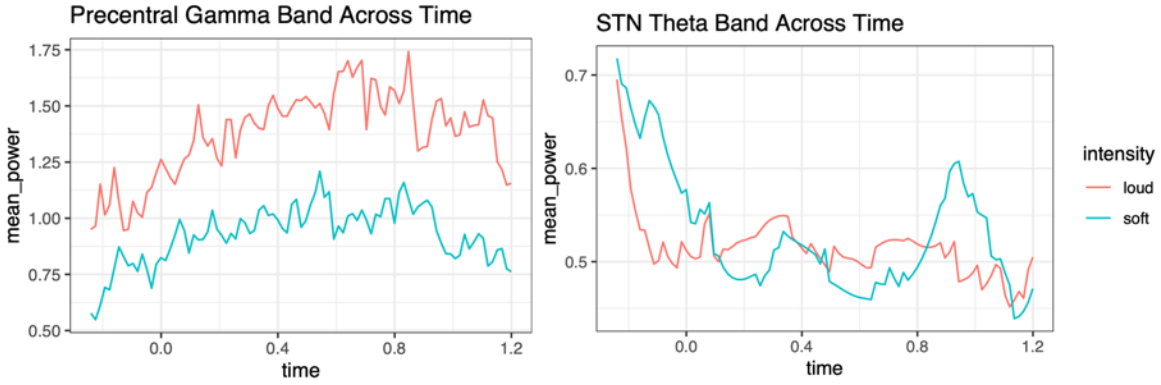


Figure 19. Temporal Differences in Vocal Intensity Encoding

Differences in the raw power were visualized between the loud and soft speaking conditions. Neural data was averaged from the electrodes demonstrating a positive correlation to

vocal intensity in each region. Trials were spilt between loud and soft productions using the median value, averaged across trials and visualized across time.

7.4 Discussion

These results present the first direct electrophysiologic evidence of vocal intensity production encoding in the basal-ganglia cortical loop. Observed effects in both the STN and precentral gyrus agree with broad findings that vocal intensity encoding requires a diffuse network across multiple neurological nodes (Baum & Pell, 1999; Liotti et al., 2003; Narayana et al., 2010). Importantly, these findings emphasize the subcortical influence of vocal intensity production, which has received only limited attention in prosody research (Paulmann, Pell, & Kotz, 2008; Pichon & Kell, 2013). Vocal intensity encoding within gamma and theta bands reinforces current hypotheses that gamma and theta band are critical neural oscillations for speech production (Giraud & Poeppel, 2012; Hyafil et al., 2015). Cumulative evidence suggests a gamma-theta vocal intensity production network including the precentral gyrus and subthalamic nucleus.

Spatially Selective Regions in Precentral Gyrus Enhance Vocal Intensity Output

Indirect imaging studies of prosodic encoding have identified that many cortical structures are involved prosodic manipulations (Baum & Pell, 1999). The left hemisphere is often relegated to 'linguistic prosody' control, yet the specific contributions remained unclear. Averaged data presented here show that the precentral gyrus gamma band power negatively predicts vocal intensity during the reaction time interval, indicating that precentral gamma band power desynchronization may contribute to motor preparation for vocal intensity encoding. This is in

contrast to some reports of phasic increases in gamma band activity prior to movement (Tzagarakis, West, & Pellizzer, 2015), but this may reflect a unique gamma dynamic for speech production. Pre-movement gamma activity is often attributed to movement preparation (Schoffelen, Poort, Oostenveld, & Fries, 2011), which may also apply to vocal intensity encoding.

Findings from the averaged data during speech production demonstrated that increases in precentral gamma band power are associated with decreases in vocal intensity production. Although initially counterintuitive, these findings of decreased activity with increasing vocal intensity in precentral gamma band agree with recent LSVT® outcomes. Children with cerebral palsy who increased vocal intensity after a course of LSVT® were found to have a pre-post decrease in BOLD signal in motor areas, including the precentral gyrus (Bakhtiari et al., 2017). The authors hypothesized that this decrease indicated less reliance on feedforward commands for vocal intensity generation; however, the results of the present study replicated this effect without a course of therapy, indicating that decreases in cortical motor gamma band may be a normal neural pattern for intensity modulation. These two studies show that the average precentral gyrus activity decreases with increasing vocal intensity, measured across two different age groups and disease etiologies and the neural activity in the present study was not attributable to disease severity.

Topographical assessment of these cortical effects reveals a more nuanced role for motor control of vocal intensity. Most recording locations in the precentral gyrus displayed a negative correlation with vocal intensity, with higher gamma band power associated with lower vocal intensity. Although gamma band has been associated with increases in movement (Nowak, Zich, & Stagg, 2018), its role in movement has been variable with negative associations with movement parameters (N. J. Ray et al., 2012). It is possible that decreases in vocal intensity

correspond to motor effector regions which interfere with movement necessary for vocal intensity increases, and therefore, require some amount of suppression. Transcranial stimulation of mouth regions within the precentral gyrus has resulted in increases in vocal intensity production (Dias et al., 2006). It is possible that the four locations which positively correlated with vocal intensity were responsible for mouth movements, similar to the study by Dias and colleagues. Another explanation may be that these positively associated regions are responsible for respiratory control, which is the primary drive of vocal intensity production (Finnegan et al., 2000). The decreases in gamma band power in most recording locations may reflect suppression of competing motor programs, allowing the selected intensity generating regions to dominate. Recent research has found that precentral gyrus gamma band corticomuscular coherence occurs in a spatially selective manner during wrist movement, with coherence only occurring in the motor regions involved in the task (Schoffelen et al., 2011), lending credence to the hypothesis that a small, spatially selective group of locations may encode for vocal intensity in the precentral gyrus.

In addition to gamma band, theta band power was positively associated with increases in vocal intensity in the averaged model; however, this effect was not replicated in the model which assessed productions on a syllable level basis. Theta band activity has been hypothesized to parse syllable level units in speech perception (Arnal & Giraud, 2012; Hyafil et al., 2015) and speech production (Giraud et al., 2007; Giraud & Poeppel, 2012; Yuchen, Xiao, & Qingfang, 2020). The positive association with theta band power from the trial-level average of neural productions may have reflected syllable level parsing effects; therefore, accounting for the syllable level condition in the second model may have better controlled for these syllable effects, eliminating the effects.

Subthalamic Nucleus Excitation Enhances Vocal Intensity Output

Speech researchers have postulated that the basal ganglia are responsible for vocal intensity production due to indirect evidence of vocal intensity deficits in PD (Murdoch, 2001); therefore, these results confirm that hypothesis through direct evidence that the basal ganglia, specifically the STN theta band power, positively predicts vocal intensity production. The finding that the basal ganglia contributes to active vocal intensity modulation control is important given recent clinical assertions that interventions which increase vocal intensity production in patients with PD may rely more heavily on the cortico-spinal system (i.e. pyramidal system) (Behrman et al., 2020; Boutsen et al., 2018; Levitt, 2014). These results show that vocal intensity modulation in patients with PD is still modulated within the extrapyramidal system, although more research is needed to determine the complete network.

Not only does vocal intensity modulation occur within the extrapyramidal system, but the STN appears to play a primary role in this process, evidenced by the stark contrast between the number of recording locations positively encoding for vocal intensity in the STN (35) compared to only five in the precentral gyrus. This difference is even more apparent considering that the number of total recordings locations in the precentral gyrus was 373 compared to only 111 in the STN. Current models of speech motor control attribute motor program sequencing and initiation to the basal ganglia-cortical loop (Bohland et al., 2010; Guenther, 2015) which is in line with some theories of basal ganglia function; however, these new data suggest that the basal ganglia may also contribute to the modulation of motor commands, specifically for vocal intensity. Results also agree with recently reported findings that articulatory gain is modulated within STN theta and

alpha bands (Dastolfo-Hromack et al., 2021). Together, these findings suggest that the STN may contribute to generalized movement gain modulation for speech production.

These results are also relevant to explain the role of the STN in movement. Much of the STN literature on movement scaling focuses on binary aspects of movement (i.e. go/no go responses) (Androulidakis et al., 2007; Cassidy et al., 2002; Kuhn et al., 2004), whereas vocal intensity encoding is a nonbinary range of values only limited by the individuals' physiologic endpoints. Descriptive results of the neural activity in locations positively predicting vocal intensity demonstrate a fluctuating pattern throughout the course of the task, which contrasts with the sustained activity of the precentral gyrus. These results imply that the STN's role in movement is much more complex than binary release of motor commands and requires more research in both speech and generalized motor systems. It is possible that some of this time-varying fluctuation in the current data is due to the complex nature of speech production. A three-way interaction was present in the STN, indicating that the phoneme produced influences the vocal intensity output. It is possible that this time-course is due to differences in phoneme production throughout the task, which may be addressed in future studies.

Complementary Roles of Gamma and Theta in Vocal Intensity Encoding

It has been suggested that gamma and theta bands are critical oscillations for the neural control of speech production and perception (Giraud et al., 2007; Giraud & Poeppel, 2012; Sengupta & Nasir, 2016) and the presented results support that prediction. Giraud and colleagues posited that gamma band and theta band are two critical oscillations which encode the temporal structure of speech acoustics. The slow fluctuation of theta band (4-8 Hz) aligns with slow-changing aspects of the speech signal, such as vocal intensity prosody manipulations, while gamma

aligns with faster temporal events. Results of theta band positively encoding vocal intensity across multiple locations in the STN support the concept that theta band encodes slow temporal events, like prosodic manipulations. Theta band activity is thought to encode syllable level structure, which matches the structure of the speech productions in the present study.

In addition to time-scale events, gamma and theta band have been associated with speech motor adaptability, specifically during the task of formant frequency adaptation learning (Sengupta & Nasir, 2016). Authors noted that high frequency oscillations (i.e. gamma band) couple with low frequency oscillations (theta or beta band) across brain networks to produce a change in the speech signal. Although the presented study did not address coherence between gamma and theta bands, the mirror image of a widespread negative correlation in precentral gamma with widespread positive correlation in the STN (Figure 17) is suggestive of coordinated processing between the STN and precentral gyrus. Gamma band has been implicated in cortical-STN coupling, with an increase in synchrony between STN single unit firing and gamma band power during movements with fast reaction times (Fischer et al., 2020). The STN is known to inhibit the motor cortex via the thalamus, and the observed increase in power with increasing intensity may relate to the inhibition observed in the precentral gyrus. Cross frequency phase coherence is the mechanism proposed by Sengupta and colleagues for motor speech adaptation and would a fruitful topic to address in future studies. Cross-frequency coupling has previously been identified during speech production (Sengupta and nasir 2015).

Gamma band is largely considered to be prokinetic and related to focal functional anatomy (Brittain & Brown, 2014; Crone et al., 1998; Pfurtscheller et al., 1993); therefore, it surprising that a decrease in gamma band power was associated with an increase in vocal intensity. However, some recent evidence demonstrates that increases in gamma band are associated with decreased

amplitude of movement (Fischer et al., 2017), which is opposite of the effect one would expect if gamma is prokinetic and aligns with results from this study. The role of gamma band is still under investigation, and these data show that gamma band may not enhance all aspects of movement across all regions. The presence of four regions in the precentral gyrus which positively encoded for vocal intensity illustrates the focal nature of gamma band oscillations in the precentral gyrus, which corresponds to neuron firing rate (Nir et al., 2007). Gamma and theta were shown to contribute to feedforward commands for formant frequency adaptations (Sengupta & Nasir, 2016), and these results demonstrate another speech acoustic variable, vocal intensity, which may be modulated by these same bands. Cumulatively, these data suggest that gamma and theta band modulate vocal intensity in a complementary fashion. It is possible that theta encodes broad movements goals, such as effort or scaling (Anzak et al., 2012; Mazzoni, Hristova, & Krakauer, 2007) which relate to vocal intensity, while gamma band encodes focal specificity of articulator movements for vocal intensity, consistent with prior literature on other speech movements.

Conclusions

Vocal intensity production is encoded within the basal ganglia-cortical loop via gamma and theta band power in the precentral gyrus and subthalamic nucleus. One of the goals of the project was to determine which region contributed greater to vocal intensity encoding, and these data demonstrate that neither contributes more, rather, each region contributes differently. Topographical effect patterns suggest that cortical encoding of vocal intensity is spatially selective, with many regions exhibiting reduced power, while more anatomically relevant region increase power. Furthermore, STN activity broadly enhanced vocal intensity production across most sites, indicating a generalized role for amplitude modulation in the basal ganglia. These differences

should be investigated in future studies and likely indicate that the precentral gyrus and STN participate in a vocal intensity network.

Limitations

These effects were found in patients with Parkinson Disease, and although the analysis of the UPDRS score and time since diagnosis concluded that disease severity did not predict our outcome measures, influence from the disease process is still possible. Geometrical differences between the cortex and STN may also contribute to differences in findings.

8.0 Synthesis and Future Direction

Data from this work substantiates and characterizes the neural basis of *phonetic-intensity* encoding, which was originally identified by behaviorists, both theoretically (Keating, 2003; Keating & Shattuck-Hufnagel, 2002) and therapeutically (Dromey et al., 1995; L. O. Ramig et al., 1996; Sapir et al., 2007). The existence of this effect has implications for speech production theory and informs current clinical procedures. This section addresses these two topics in greater detail than described in the manuscript section. The section culminates in potential future directions of this work.

Theoretical Implications

Earlier sections described the potential impact of this work on multiple speech production theories (Section 2.2). As described earlier, this work incorporated components of prominent speech motor control perspectives: 1) the oscillatory nature of Dynamical Systems Theory (DST) and 2) the neural locus of vocal intensity control and importance of acoustic information in the internal model. The finding of *phonetic-intensity encoding* in theta band activity and vocal intensity encoding in both gamma and theta bands within the cortico-basal ganglia loop has implications for these theories. Results also provide preliminary answers to the hypotheses in question, namely: 1) Aim 1: Do populations of neurons exist in the precentral gyrus and subthalamic nucleus which control both movements to generate phoneme patterns and vocal intensity increases (*phonetic-intensity encoding*)? and 2) Aim 2: What are the respective contributions of the STN and precentral gyrus in the control of vocal intensity output?

These results provided some direct evidence to answer the immediate experimental questions. To answer the first question, populations of neurons do exist in the subthalamic nucleus

that encode for *both* phoneme and vocal intensity production, which supports the hypothesis that *phonetic-intensity encoding* is neurologically mediated. However, the second portion of this hypothesis was not found; the precentral gyrus encodes both parameters, but not in single neuron populations. Phoneme encoding and vocal intensity encoding were identified in separate recording locations. To answer the second question, results from the second aim showed that the STN and precentral gyrus both contribute to vocal intensity production, but the patterns were different. Averaged activity in the precentral gyrus gamma band decreased while vocal intensity increased, and a small portion of locations within the precentral gyrus positively correlated to vocal intensity. This pattern may indicate a spatially selective enhancement of vocal intensity in the precentral gyrus which depends on reduction of competing motor programs and enhancement of articulator regions which facilitate intensity. Alternatively, theta band activity during speech production robustly demonstrated numerous locations in the STN which positively encoded for vocal intensity. This indicates a generalized role for speech amplitude modulation for the subthalamic nucleus. These results have implications for speech motor control theory as well clinical STN-DBS outcomes.

The presented results validate the assertions of Keating and colleagues that the behavioral observation of *phonetic-encoding* of prosodic structure needs to be modeled as a fundamental neurological control mechanism for speech (Cho & Keating, 2001; Keating, 2003; Keating & Shattuck-Hufnagel, 2002). The finding of multiple neurological sites encoding both phoneme and intensity information brings to light the fundamental neurological basis of *phonetic-intensity encoding* and not purely biomechanical or behavioral processes. Keating's original discussion highlighted that Levelt's model of speech production (Levelt et al., 1999) did not explicitly model a link between the segmental phonetic features and prosodic structure (Keating, 2003; Keating &

Shattuck-Hufnagel, 2002). Per Levelt, stored syllable plans are the basic unit for phonetic encoding, and gestures are encoded within this syllabic plan (Levelt et al., 1999); however, Keating argues that even if these plans are stored, the retrieval of this plan requires subsequent processing and is highly sensitive to prosodic features (Keating, 2006). Vocal intensity is one of the components of prosody, and the finding of overlap between intensity and phoneme production supports the concept that if stored plans are activated, they may be further processed to include information about vocal intensity. These conclusions are strengthened by the finding that spectral and temporal changes occurred as patients increased vocal intensity, indicating an alteration to the segmental/phonetic gestures. The data from this study did not include the other aspects of prosody (duration, pitch), and did not analyze multi-syllabic prosodic structure; therefore, more information would be required to make a definitive statement on prosody as a whole and explicit alterations to Levelt's model. However, these results offer preliminary evidence that an explicit link may exist between the planning of prosody and segmental features, and this may occur within the basal ganglia. DST and the internal model also provide some incomplete explanations for the encoding of prosody and segmental features. Results are interpreted within the most relevant portions of each of these theories.

Dynamical Systems Theory

Dynamical systems theory literature focuses on the oscillating nature of speech production and proposes that different articulatory and prosodic elements overlay on one another in a cyclical fashion (Barbosa, 2007; Liss et al., 2010; Tilsen, 2009) (See Figure 2). This theoretical perspective aligns well with the concept of neural oscillations, which were investigated in both theta and gamma bands. The finding that both features (phoneme and vocal intensity) were encoded within

the same oscillatory band (theta band) agrees with the concept of speech gesture and prosody overlapping in DST. According to DST, oscillatory movements for phonemic gestures are nested within a higher level of prosody generation. Referencing Figure 2 in section 2.2.2, gestural-phonemic movements (i.e. "spa") are cyclically embedded within the higher levels of prosodic involvement. Specifically, the phonemic movements are proposed to be encoded within the same oscillation frequency as the syllable-level prosodic feature (labeled as CVC in Figure 2). The presented study focused on syllable-level production features, so the finding of theta band encoding both phoneme and vocal intensity production agrees with this prediction. The study did not investigate higher levels of prosodic function (phrase, sentence, discourse etc.), which may further inform these patterns.

Findings of gamma band encoding vocal intensity in the precentral gyrus does not agree with the proposed oscillation structure in DST. Gamma band is a high-frequency oscillation, and prosody is proposed to occur in a relatively slow frequency to encompass larger speech segments (Tilsen, 2009). If *phonetic-intensity encoding* occurs within theta band as found in the STN, then it would be expected that prosodic features, such as intensity, would be encoded within theta band or slower oscillations, not high frequency oscillations like gamma. It is also possible that this relative, hierarchical oscillation structure (slow frequencies for prosody, fast frequencies for articulation) is unique within neural regions. For example, theta band encodes both phoneme and vocal intensity within the STN, but it may be that gamma band encodes both features within the precentral gyrus. A single electrode location encoding both vocal intensity and phoneme production was *not* found within the precentral gyrus, but both features were encoded in separate locations within precentral gamma band. With this interpretation, the final piece to understand would be how intensity encoding in the STN (theta band) interacts with vocal intensity encoding

in the precentral gyrus (gamma band). The theory does not currently describe an interpretation for the same feature (i.e. syllable level prosody) being encoded in multiple oscillation levels.

Behaviorally, the presented results replicated the finding that acoustic adjustments are made as participants increase vocal intensity (Tjaden & Martel-Sauvageau, 2017), which can be interpreted as a behavioral 'attractor' state like the findings of vocal fold register shifts (Steinhauer, 2000), generally supporting DST applications for speech production. The question then remains if this 'attractor state' occurs because of overlapping neural oscillatory behaviors. Evidence of *phonetic-intensity encoding* was found in STN theta band, showing that some overlap does exist within a single oscillation band. On a surface level, this finding supports DST as mentioned earlier. However, it should be noted that although *phonetic-intensity encoding* was observed in theta band, phonetic encoding alone was not observed. A main effect of phoneme (i.e. phonetic encoding) was observed in gamma band, a higher oscillating frequency, and in the precentral gyrus. This may indicate that vocal intensity is the primary feature encoded within theta band, and it encodes this feature in a phonetic topography, which is distinct from primarily encoding phoneme production as a main effect. Prior findings of articulatory gain were not identified in gamma band, but rather low frequencies of theta and alpha band (Dastolfo-Hromack et al., 2021), indicating that 'gain' may be a low frequency phenomenon. A segregated architecture is found in primate studies of the basal ganglia showing that the basal ganglia encode movement parameters in specialized regions according to movement effector and neural region (i.e. segregated loop) (Alexander et al., 1986; M. DeLong & Wichmann, 2010). These results for *phonetic-intensity encoding* may therefore suggest that low oscillatory frequencies encode vocal intensity in a segregated architecture as in primate limb studies. To further disentangle these findings, more research is

needed examining *phonetic-prosodic encoding* in multiple basal ganglia nuclei and anatomically connected structures.

Internal Model

Findings of both vocal intensity and phonetic intensity encoding have implications for the internal model of speech production as well. Instantiations of the internal model of speech production (DIVA and GODIVA models) propose that portions of the basal ganglia (globus pallidus, putamen, substantia nigra pars reticulata) are involved in phoneme/motor plan selection as well as motor program initiation (labeled as "Go Signal") (Bohland et al., 2010; Guenther, 2015). These predictions are in line with basal ganglia functions but do not encompass all the known basal ganglia control parameters such as amplitude scaling (Anzak et al., 2012; Tan, Pogosyan, Anzak, Ashkan, et al., 2013). The finding of *phonetic-intensity encoding* in addition to *intensity encoding* in the STN indicates that the role of the basal ganglia may be expanded to include motor program modulation for vocal intensity. Although gain functions are not explicitly rejected as a basal ganglia parameter in this model, it is suggested that motor cortical areas (i.e. areas that are responsible for speech articulation) are responsible for prosody production, "prosody generator" location (Guenther, 2015). This interpretation is in line with dominant literature in prosody, which focuses on cortical control; however, this new subcortical evidence supports an alternate theory that vocal intensity/prosody production may be sub-cortically driven (Baum & Pell, 1999; Paulmann et al., 2008; Pichon & Kell, 2013). Finally, the STN is not explicitly noted within the model as an important basal ganglia node for speech production, but this interpretation may be revised given the current evidence.

Within the DIVA model of speech production, segmental features (i.e. phonemes) are considered to be encoded differently than suprasegmental features (i.e. prosody/intensity) primarily because of the differential response to auditory deprivation. However, these results indicate that although phoneme information and vocal intensity information may respond differently, the motor parameters co-modulate within the basal ganglia. Per the DIVA model, the 'prosody generator' is located within the same cortical regions which control language processing (Guenther, 2015). For vocal intensity, this appears to be partially true. Five locations in the precentral gyrus encoded positively for vocal intensity production; however, many more locations in the basal ganglia encoded for vocal intensity. This may indicate that the prosody generator is not exclusively located in motor cortical regions but may be subcortical as well. Furthermore, there are no explicit links described between the prosody generator and the speech sound map or articulator map in the pre-movement planning, only in online processing, which accounts for reflexive and non-reflexive alterations to the motor plan which is already being produced (Guenther, 2015; R. Patel et al., 2011; R. Patel et al., 2015). Considering that vocal intensity parameters are selected before the movement begins, the model should reflect this aspect in selection and initiation, along with phonemic elements. This study only examined one aspect of prosody, vocal intensity, and more research would be needed in the other aspects of prosody production to definitively change the location of the 'prosody generator,' as well as more timing studies to assess the extent to which vocal intensity and phoneme are encoding during motor planning stages, but these results do highlight an important area for future research and discussion.

Theoretical Summary

To summarize the theoretical implications, these findings partially support concepts and predictions of DST and the internal model; however, specific aspects of the theories need to be addressed with more research, and potential alterations. First, DST behavioral studies indicated that oscillations occur in hierarchical levels, with prosody occurring at the slowest time-interval, yet these results demonstrate vocal intensity modulations in a high frequency oscillation. The oscillation patterns, therefore, may be relative to neural region and process and not absolute hierarchies. Second, the internal model of speech production, as described by the DIVA model indicates that the prosody generator is in the cortical areas for speech motor control; however, multiple vocal intensity-dependent neural locations were found which implicates the STN in this process. Therefore, the proposed location of the prosody generator may be expanded to the basal ganglia, provided that these results are replicated across pitch and duration components of prosody.

Clinical Implications

This research study was not intended to directly inform clinical outcomes; however, the data gathered may inform clinical findings in STN-DBS, which have been found to be highly variable and dependent on the speech measurement (Aldridge et al., 2016). Some studies have found that STN-DBS improves vocal intensity production (Dromey et al., 2000; Gentil et al., 2001), while others see no effect (Tripoliti et al., 2008). Conversely, speech articulation and intelligibility have seen more consistent results with most studies findings that articulation worsens in a subset of patients (Hammer, Barlow, Lyons, & Pahwa, 2011; Pinto et al., 2014; Tripoliti et al., 2014; Tripoliti, Zrinzo, et al., 2011). It has been proposed that that the negative outcomes on speech articulation may arise from stimulation affecting areas outside of the STN, and invading other tracts, such as cerebello-thalamic (Astrom et al., 2010; Tanaka et al., 2015; Tripoliti et al.,

2008). Although this study did not directly assess tracts, it does inform our knowledge of the potential speech functions located directly within the STN which could be affected. Because *phonetic-intensity encoding* was identified within the STN, it is possible that stimulation in these regions could be directly, rather than indirectly, affecting speech articulation.

Another hypothesis generated from STN-DBS studies is that the frequency of stimulation impacts articulatory outcomes. Tornqvist and colleagues found that high frequency stimulation in the STN (>130 Hz) deteriorates articulatory precision (Tornqvist, Schalen, & Rehncrona, 2005). Gamma band tested in this study had a range of 70-150 Hz, which is not an exact comparison; however, it is informative that neither of the speech functions (phoneme or intensity) were associated with gamma band function in the STN. It is possible that high frequency oscillations encode non-speech functions, such as limb motor control. Low frequency stimulation may enhance speech functions within the STN due to the finding of speech parameters being encoded in theta band, while high frequency stimulation interferes with typical speech processing. Low frequency has been found to improve intelligibility in patients with STN-DBS (Grover et al., 2019). However, more research is needed on this topic as this frequency stimulation parameter was only manipulated in two studies.

Limitations

Although this study yielded valuable information about the theoretical and clinical role for the basal-ganglia-cortical loop in speech production, limitations were present which restricts the scope of interpretation.

- *Articulator Identification*: Four consonants were included in the design of the experiment because these consonants have previously been identified in the motor

cortical areas; however, the variability of articulator involvement within these phonemes was not ideal. For determining functional articulator mapping, a more articulator-diverse set of consonants would have been helpful to identify neural regions encoding a specific articulator. This low variability limits the extent to which articulator effects can be interpreted and opens the possibility that given more variability in phonemes, neural articulator positions may have been different.

- *Parkinson's Disease:* This research was only possible because patients were undergoing a routine clinical procedure to treat Parkinson's disease. Although disease severity measures were included in the modeling procedures to control for this aspect, some neural physiology differences are likely present which limits the ability to generalize these results to normal populations.
- *Operating Room Conditions:* Measurements were completed during normal clinical operations, which includes background noise from clinicians speaking and clinical machines running. Although the room was kept as quiet as possible during the actual recording procedures and acoustic measures were screened for extreme background noise, this noise may have affected results. Patients with Parkinson's disease respond to the Lombard effect (Heracleous et al., 2012), during which vocal intensity increased reflexively due to background noise. Therefore, varying levels background noise could reflect this process rather than completely volitional control.
- *Vocal Intensity Perception Control:* Some neural processing may reflect patients hearing and perceiving their own voice. The motor theory of speech perception indicates that motor cortical areas are activated during speech perception (Liberman

& Mattingly, 1985). Due to time constraints within the operating room/clinical setting, no data were collected while patients were strictly listening and not listening with the intent to produce speech sounds. A pure perception-based data collection element may have informed this question.

- *Vowel Variability*: The design of the experiment operationalized 'phoneme' as the consonant, even though the entire syllable, including the vowel, was measured for vocal intensity. The syllable position and frequency of the phonemes, both vowel and consonant, were controlled for in the stimuli; however, the inherent variability in vowel production for each of the syllables was not controlled for in the modeling and likely introduced extra neural variability. The practice of utilizing the initial phoneme (consonant) as the most salient feature, despite variability in the remainder of the utterance, has been utilized in prior work (Chrabaszcz et al., 2019).

Future Directions

The findings of *phonetic-intensity encoding* and *intensity encoding* within the basal-ganglia cortical loop are informative to both theory and clinic; however, more research is needed. First, theories on speech motor control, as discussed earlier, typically do not focus on vocal intensity as a sole parameter, but rather prosody as a generalized speech concept. Prosody is a much more complex phenomenon than just vocal intensity and encompasses duration and pitch as well as syllabic structure. More research is needed which examines the basal ganglia's role in pitch and duration processing of speech, both individually and as an integrated prosodic unit. Furthermore, cognitively-based studies on prosody identify two different kinds of prosody, emotional and

linguistic. This study did not attempt to disentangle these components, but rather controlled for it by completing intensity modulations in a non-speech task. Only after this type of research would complete alterations to the origins of prosody generation be appropriate in speech production models.

Further disentangling the specific articulators in the precentral gyrus which are positively associated with vocal intensity would also be highly informative. In earlier sections, it was discussed that these positively encoding regions may subserve respiratory function, which drive vocal intensity changes. However, respiratory measurements were not obtained during this experimental procedure. Research investigating the relative neural contributions from specific articulators to vocal intensity control would be highly informative.

The data, particularly from the vocal intensity findings, suggest that the STN and cortex are engaging in a systematic network. This hypothesis should also be tested. Theories indicate that theta and gamma band interact, which can be observed in cross-frequency coupling measures like coherence (Lizarazu et al., 2019; Sengupta & Nasir, 2015). Understanding if and how these two signals are related to one another would be valuable to understanding speech production network dynamics and illuminate the broader processes involved in intensity encoding.

As discussed earlier, these results also inform clinical procedures and may indicate differential roles for high vs. low frequency activity in the subthalamic nucleus. It would be interesting and informative to repeat this vocal intensity increase procedure, but in the office ON and OFF stimulation while varying frequency of stimulation parameters. Testing the hypothesis that high frequency stimulation degrades vocal intensity and articulation parameters and low frequency (theta band range) stimulation would improve these parameters would be highly valuable. Results could substantiate the claims made here regarding the role of low vs. high

frequency stimulation, as well as inform clinical procedures. Additionally, it would be interesting to know if these *phonetic-intensity encoding* behavioral effects can be replicated under different stimulation parameters.

Summary

These data provide initial neurological evidence of *phonetic-intensity encoding* and *vocal intensity encoding* in the STN and precentral gyrus. The presence of this effect informs speech motor control theories, particularly DST and the internal model, which incompletely describe this phenomenon. Widespread findings of intensity and *phonetic-intensity encoding* in the STN provides support for the notion that prosody may not be completely cortically driven, and that subcortical processing of speech function does, in fact, occur. Finally, effects were found in both high and low frequency bands, which informs interpretations for clinical STN-DBS findings and provides an oscillation-based framework for understanding vocal intensity production.

Appendix A Experimental Procedures

EXPERIMENTAL SET-UP

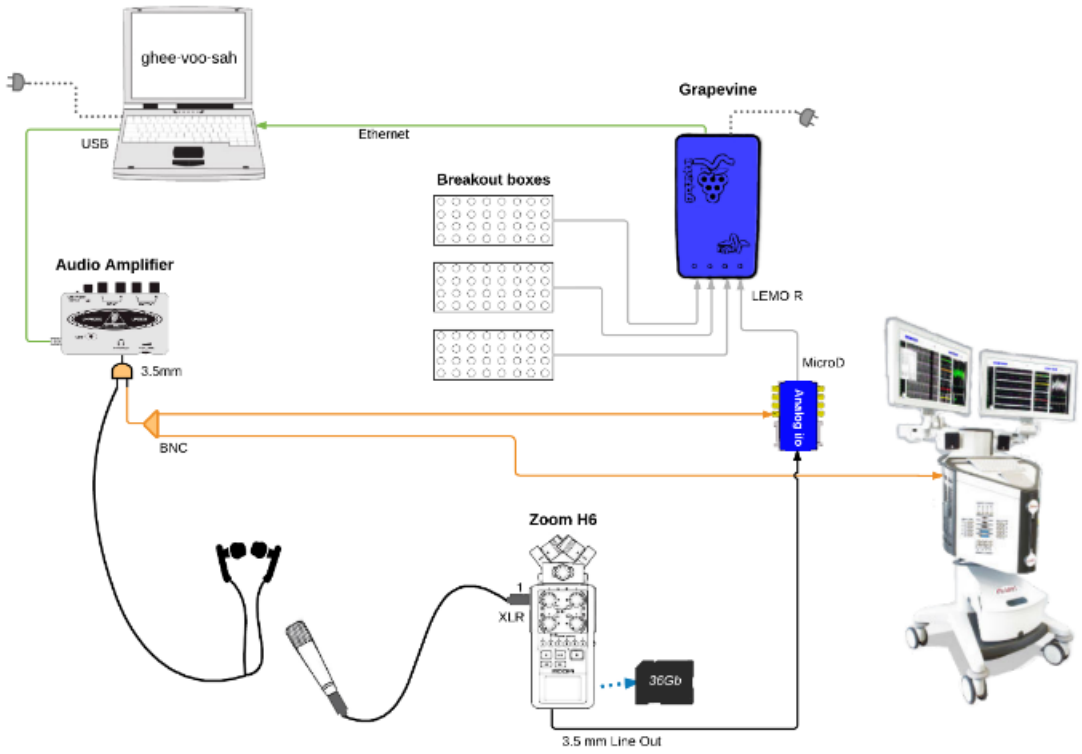
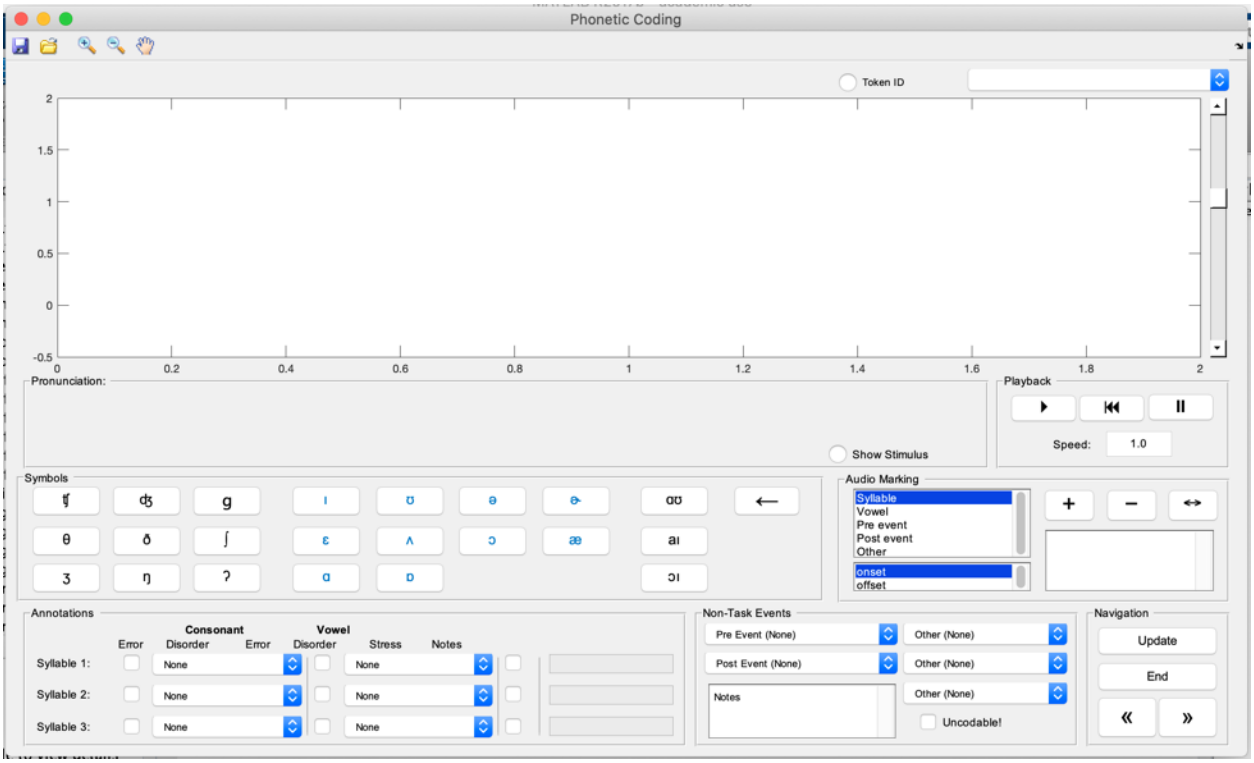


Image courtesy of Brain Modulation Lab

CUSTOM GUI UTILIZED FOR PHONETIC CODING



GUIDELINES FOR PHONETIC CODING

The following instructions were given to the trained speech-language-pathology students completing transcription on the behavioral data. These instructions served as a guideline and anchor for coding procedures across coders and were accompanied by a training interval as well as regular follow-up and feedback on coding processes.

Before you start coding, be in a quiet place and always use provided headphones.

For each word, make mark the timing by placing the cursor for each of the following:

1. Onset of each syllable
   Warning [01] if number of syllable onsets is different from 3, except if no phonetic code is present in which case no syllable onset is expected.
   Warning [02] if phonetic coding is missing and syllable onsets are given.

2. Offset of each syllable (DO NOT mark this if the acoustics are continuous between phonemes)
   Warning [03] if syllable offsets are present but no syllable onsets are given
   Warning [04] if last syllable has no offset
   Warning [05] if some syllable has more than one offset, i.e. more than one syllable offset between consecutive syllable onsets.

3. Onset of vowel

Guidelines for marking timing on error productions:

4. If there is no consonant at the beginning of a syllable (only a vowel), mark the beginning of the vowel as a syllable onset, and do not mark it as a vowel onset. The offset in this case will still be a syllable offset (not vowel).
   Warning [06] if more than one vowel onset in syllable
   Warning [07] if vowel onsets are given but not the syllable onsets
   Warning [08] if vowel onset falls outside syllable.

5. Vowel offsets are only marked in cases of error, when the patients said a CVC syllable, instead of a CV syllable. In this case you would need a vowel offset, AND a syllable offset, and the phonemes should be CVC/CV/CV (for example).
   Warning [09] if more than one vowel offset in syllable
   Warning [10] if vowel offsets are given but not the syllable onsets
   Warning [11] if vowel offset falls outside syllable.

6. If a patient made a sound that is acoustically connected to the onset or offset of the target production, mark it as a 'pre-event' or 'post-event'. When marking a pre or post event, it is assumed that the offset of the 'pre' event is the same as the onset of the first syllable, and the onset of the 'post' even is the same as the offset of the final syllable (therefore acoustically connected). Therefore, coders only code the onset of the pre-event and an offset of a post event.
   Warning [12] if more than one Pre onset time is given
   Warning [13] if pre event onset is not before first syllable onset
   Warning [14] if one or more pre event offsets are present

Warning [15] if one or more post event onset(s) are present
Warning [16] if more than one Post event offsets are given
Warning [17] if post event offset is not after 3rd syllable offset.

7. Sound unrelated to the target production (coughing, background noise etc.) should be indicated with the appropriate label and marked with its own timing using the cursor tool.

Qualitative categorization of productions:

Please use the appropriate drop-down box to label any pathological characteristics that you hear. Use the terms below:

- Distortion: Production is so indistinct that the intended phoneme is unclear.
- Imprecise: Speech sound is 'underarticulated', but the intended phoneme can be clearly identified.
- Spirantization: Visible or audible frication noise occurs in the plosive.
- Dysfluency: Repetition or prolongation of speech sound.
- Voice Quality Terms:
- Creaky voice
    - Tremor
    - Breathy voice
    - Hoarse/Harsh
    - Voice Break
    - Strain

General guidelines for accurate marking of speech sound timing:

Fricatives:

1. Use the overall energy (broad spectrum energy) as a guideline for marking onset and offset
2. Voiced fricatives will have some voicing (glottal pulse) in the beginning of the sound, please include that in marking.

Vowels:

1. Always looking for the onset of voicing (vertical glottal pulses).
2. If preceded by a voiced consonant, look for initiation of a formant structure.
3. Be aware that diphthongs will have formants for two vowels with the vowel structure

Nasals:

1. Look for two clues to identify the start of nasal energy:
    - Abrupt reduction in overall energy
    - Second formant shift
2. When marking offset of final position nasals, mark the final glottal pulse as the end of the word

WORD LISTS

The following pages include the word lists utilized for the experiment. These triplet phrases were recorded and played back to the participants via earbuds, and they were instructed to repeat these nonsense syllables.

WORD LIST #1

1	'tee'	'soo'	'vah'
2	'SAH'	'TEE'	'GHOO'
3	'ghah'	'tee'	'voo'
4	'TEE'	'GHOO'	'VOO'
5	'SEE'	'GHAH'	'TOO'
6	'TAH'	'VOO'	'SEE'
7	'TOO'	'GHEE'	'VAH'
8	'vah'	'see'	'tah'
9	'vee'	'tah'	'ghoo'
10	'TAH'	'SEE'	'VEE'
11	'TOO'	'SAH'	'TEE'
12	'tah'	'see'	'vee'
13	'ghee'	'voo'	'sah'
14	'GHAH'	'VEE'	'SOO'
15	'TEE'	'VOO'	'SOO'
16	'VOO'	'SAH'	'TOO'
17	'ghoo'	'see'	'tah'
18	'GHAH'	'TEE'	'VOO'
19	'VOO'	'TAH'	'GHEE'
20	'ghoo'	'tah'	'see'
21	'sah'	'vah'	'ghoo'
22	'VEE'	'GHAH'	'SOO'
23	'SAH'	'VAH'	'GHOO'
24	'VAH'	'SEE'	'TAH'
25	'SOO'	'GHEE'	'TAH'
26	'VAH'	'SOO'	'GHAH'
27	'soo'	'voo'	'vee'
28	'soo'	'vah'	'ghee'
29	'GHOO'	'VEE'	'TAH'
30	'too'	'ghee'	'vah'
31	'SOO'	'TEE'	'GHAH'
32	'see'	'soo'	'ghah'
33	'GHEE'	'VOO'	'SAH'
34	'vah'	'too'	'tee'
35	'GHEE'	'SAH'	'TOO'
36	'vah'	'ghoo'	'tee'
37	'VAH'	'GHOO'	'TEE'
38	'tah'	'soo'	'vee'
39	'sah'	'ghee'	'voo'
40	'VEE'	'TAH'	'GHOO'
41	'SAH'	'VEE'	'GHOO'
42	'VEE'	'GHOO'	'SAH'
43	'VAH'	'TOO'	'TEE'
44	'SOO'	'TAH'	'GHEE'
45	'GHOO'	'SEE'	'TAH'
46	'too'	'see'	'tee'
47	'tee'	'voo'	'soo'
48	'VOO'	'SEE'	'TAH'
49	'soo'	'ghee'	'tah'
50	'TOO'	'VAH'	'GHEE'
51	'VOO'	'TEE'	'SAH'
52	'sah'	'tee'	'ghoo'
53	'voo'	'tah'	'ghee'
54	'vee'	'too'	'sah'
55	'SAH'	'TOO'	'VEE'
56	'soo'	'tah'	'ghee'
57	'SOO'	'VAH'	'GHEE'
58	'sah'	'too'	'vee'
59	'voo'	'sah'	'too'
60	'voo'	'see'	'tah'
61	'vah'	'soo'	'ghah'
62	'SEE'	'VEE'	'GHOO'
63	'vee'	'ghoo'	'sah'
64	'GHEE'	'TAH'	'SOO'
65	'TAH'	'GHOO'	'VEE'
66	'see'	'ghah'	'too'
67	'too'	'vah'	'ghee'
68	'sah'	'vee'	'ghoo'
69	'TAH'	'GHEE'	'VOO'
70	'tah'	'voo'	'see'
71	'ghah'	'too'	'see'
72	'VEE'	'GHAH'	'TOO'
73	'ghee'	'vah'	'too'
74	'ghah'	'vee'	'vah'
75	'GHAH'	'SOO'	'VAH'
76	'VOO'	'TEE'	'GHAH'
77	'tee'	'ghoo'	'voo'
78	'GHEE'	'SAH'	'VOO'
79	'vah'	'ghoo'	'see'
80	'GHOO'	'SAH'	'TEE'
81	'soo'	'tee'	'ghah'
82	'ghoo'	'ghee'	'ghah'
83	'GHOO'	'TAH'	'SEE'
84	'SOO'	'VOO'	'VEE'
85	'ghee'	'tah'	'soo'
86	'vee'	'ghah'	'soo'
87	'TOO'	'GHAH'	'SEE'
88	'vee'	'ghah'	'too'
89	'SEE'	'SOO'	'GHAH'
90	'see'	'vah'	'ghee'
91	'SEE'	'VAH'	'GHEE'
92	'voo'	'tee'	'ghah'
93	'see'	'too'	'vah'
94	'SEE'	'TOO'	'VAH'
95	'TOO'	'SEE'	'TEE'
96	'voo'	'tee'	'sah'
97	'TAH'	'SOO'	'VEE'
98	'tee'	'voo'	'sah'
99	'see'	'vee'	'ghoo'
100	'tee'	'ghah'	'soo'
101	'GHEE'	'VAH'	'TOO'
102	'ghee'	'sah'	'voo'
103	'TEE'	'SOO'	'VAH'
104	'too'	'sah'	'tee'
105	'ghoo'	'vee'	'tah'
106	'GHAH'	'TOO'	'SEE'
107	'tah'	'ghoo'	'vee'
108	'VEE'	'TOO'	'SAH'
109	'GHAH'	'VEE'	'VAH'
110	'too'	'ghah'	'see'
111	'tah'	'ghee'	'voo'
112	'ghah'	'soo'	'vah'
113	'VAH'	'GHOO'	'SEE'
114	'ghee'	'sah'	'too'
115	'ghah'	'vee'	'soo'
116	'ghoo'	'sah'	'tee'
117	'GHOO'	'GHEE'	'GHAH'
118	'TEE'	'GHAH'	'SOO'
119	'SAH'	'GHEE'	'VOO'
120	'TEE'	'VOO'	'SAH'

WORD LIST #2

1	'VAH'	'TOO'	'SEE'
2	'vee'	'tah'	'soo'
3	'SOO'	'VEE'	'TAH'
4	'too'	'vee'	'ghah'
5	'GHEE'	'GHAH'	'VAH'
6	'SOO'	'VAH'	'GHEE'
7	'SEE'	'TAH'	'VOO'
8	'see'	'ghee'	'too'
9	'tah'	'soo'	'ghee'
10	'TAH'	'SEE'	'GHOO'
11	'VEE'	'TAH'	'SOO'
12	'see'	'tah'	'ghoo'
13	'SOO'	'TEE'	'GHAH'
14	'vee'	'ghah'	'too'
15	'GHOO'	'VEE'	'TAH'
16	'vee'	'sah'	'ghoo'
17	'too'	'vah'	'vee'
18	'VEE'	'GHAH'	'TOO'
19	'VOO'	'TAH'	'VEE'
20	'voo'	'ghee'	'tah'
21	'vah'	'tee'	'voo'
22	'tah'	'voo'	'ghee'
23	'tah'	'see'	'voo'
24	'SOO'	'GHAH'	'VEE'
25	'vee'	'too'	'ghee'
26	'TOO'	'VAH'	'VEE'
27	'VEE'	'TOO'	'GHEE'
28	'TAH'	'VOO'	'SEE'
29	'TEE'	'SAH'	'GHOO'
30	'soo'	'vee'	'tah'
31	'ghoo'	'vah'	'tee'
32	'SAH'	'GHOO'	'TEE'
33	'tee'	'soo'	'vah'
34	'vah'	'soo'	'ghee'
35	'sah'	'ghoo'	'vee'
36	'ghee'	'ghah'	'vah'
37	'VAH'	'TEE'	'VOO'
38	'TOO'	'GHEE'	'VAH'
39	'TAH'	'VOO'	'GHEE'
40	'SEE'	'TOO'	'GHAH'
41	'ghoo'	'sah'	'voo'
42	'soo'	'vah'	'ghee'
43	'TEE'	'GHAH'	'SOO'
44	'SEE'	'GHOO'	'TAH'
45	'ghah'	'tee'	'vee'
46	'GHOO'	'SAH'	'VOO'
47	'ghah'	'soo'	'tee'
48	'VAH'	'SOO'	'GHEE'
49	'GHOO'	'VAH'	'TEE'
50	'see'	'ghoo'	'tah'
51	'GHEE'	'VOO'	'TAH'
52	'too'	'see'	'vah'
53	'GHAH'	'TEE'	'SOO'
54	'VOO'	'GHEE'	'TAH'
55	'GHEE'	'SEE'	'TOO'
56	'TEE'	'SOO'	'GHAH'
57	'ghee'	'ghoo'	'sah'
58	'tee'	'soo'	'ghah'
59	'VOO'	'SEE'	'TOO'
60	'VEE'	'TOO'	'SAH'

61	'TOO'	'VEE'	'GHAH'
62	'TEE'	'VAH'	'SOO'
63	'see'	'too'	'ghah'
64	'VOO'	'SAH'	'VAH'
65	'sah'	'tee'	'ghoo'
66	'tah'	'see'	'ghoo'
67	'GHOO'	'VEE'	'SAH'
68	'VOO'	'VAH'	'SAH'
69	'TEE'	'SOO'	'VAH'
70	'sah'	'voo'	'see'
71	'GHEE'	'GHOO'	'SAH'
72	'GHOO'	'SAH'	'TEE'
73	'sah'	'vee'	'soo'
74	'voo'	'vah'	'sah'
75	'SEE'	'TAH'	'GHOO'
76	'ghah'	'tah'	'voo'
77	'GHAH'	'TAH'	'VOO'
78	'soo'	'ghah'	'tee'
79	'GHEE'	'VOO'	'GHAH'
80	'too'	'ghee'	'sah'
81	'SAH'	'GHOO'	'VEE'
82	'tee'	'ghah'	'soo'
83	'VAH'	'GHOO'	'SEE'
84	'SOO'	'GHAH'	'TEE'
85	'voo'	'see'	'too'
86	'GHAH'	'TOO'	'SEE'
87	'tah'	'voo'	'see'
88	'vee'	'too'	'sah'
89	'TAH'	'SOO'	'GHEE'
90	'tee'	'sah'	'ghoo'
91	'tee'	'vah'	'soo'
92	'sah'	'ghoo'	'tee'
93	'vah'	'ghoo'	'see'
94	'ghoo'	'vee'	'tah'
95	'ghah'	'too'	'see'
96	'ghoo'	'vee'	'sah'
97	'ghee'	'see'	'too'
98	'see'	'tah'	'voo'
99	'VAH'	'GHEE'	'TOO'
100	'SAH'	'TEE'	'GHOO'
101	'ghee'	'voo'	'tah'
102	'TOO'	'GHEE'	'SAH'
103	'ghee'	'voo'	'ghah'
104	'voo'	'tah'	'vee'
105	'ghoo'	'sah'	'tee'
106	'GHAH'	'TEE'	'VEE'
107	'SEE'	'GHEE'	'TOO'
108	'TAH'	'SEE'	'VOO'
109	'SAH'	'VOO'	'SEE'
110	'soo'	'tee'	'ghah'
111	'TOO'	'SEE'	'VAH'
112	'ghah'	'tee'	'soo'
113	'vah'	'ghee'	'too'
114	'GHAH'	'SOO'	'TEE'
115	'VEE'	'SAH'	'GHOO'
116	'too'	'ghee'	'vah'
117	'soo'	'ghah'	'vee'
118	'vah'	'too'	'see'
119	'voo'	'sah'	'vah'
120	'SAH'	'VEE'	'SOO'

WORD LIST #3

1	'soo'	'tee'	'ghah'
2	'TAH'	'GHOO'	'SEE'
3	'ghoo'	'sah'	'vee'
4	'voo'	'ghee'	'ghoo'
5	'vah'	'ghee'	'soo'
6	'SAH'	'VEE'	'TOO'
7	'GHOO'	'VAH'	'SEE'
8	'VOO'	'GHEE'	'GHOO'
9	'see'	'tee'	'tah'
10	'GHOO'	'VEE'	'SAH'
11	'too'	'vee'	'ghah'
12	'see'	'sah'	'tee'
13	'ghee'	'soo'	'tee'
14	'TEE'	'VAH'	'SOO'
15	'soo'	'tah'	'ghee'
16	'SEE'	'SAH'	'TEE'
17	'soo'	'tee'	'vah'
18	'TAH'	'VOO'	'VAH'
19	'GHAH'	'TOO'	'SEE'
20	'sah'	'vee'	'too'
21	'SAH'	'VOO'	'GHEE'
22	'GHAH'	'SEE'	'TOO'
23	'ghoo'	'vee'	'sah'
24	'VAH'	'GHEE'	'SOO'
25	'tah'	'voo'	'vah'
26	'SEE'	'GHAH'	'VOO'
27	'vee'	'soo'	'ghoo'
28	'vee'	'ghoo'	'tah'
29	'ghee'	'ghah'	'sah'
30	'tee'	'ghoo'	'sah'
31	'see'	'ghah'	'voo'
32	'TAH'	'VEE'	'TOO'
33	'too'	'tah'	'vee'
34	'vah'	'see'	'tee'
35	'ghoo'	'sah'	'tee'
36	'SAH'	'TOO'	'VEE'
37	'VEE'	'TAH'	'SOO'
38	'VAH'	'GHEE'	'VOO'
39	'voo'	'soo'	'ghee'
40	'VAH'	'TOO'	'SEE'
41	'VOO'	'TAH'	'SEE'
42	'VEE'	'SOO'	'GHOO'
43	'soo'	'ghah'	'vee'
44	'SOO'	'GHAH'	'VEE'
45	'ghee'	'sah'	'voo'
46	'VAH'	'SEE'	'TEE'
47	'see'	'ghah'	'vah'
48	'GHAH'	'TOO'	'VEE'
49	'TEE'	'GHOO'	'SAH'
50	'SOO'	'TAH'	'GHEE'
51	'vah'	'ghee'	'voo'
52	'vee'	'tah'	'soo'
53	'VEE'	'GHOO'	'TAH'
54	'VEE'	'TOO'	'GHAH'
55	'sah'	'voo'	'ghee'
56	'sah'	'see'	'vah'
57	'vah'	'too'	'see'
58	'VAH'	'GHEE'	'TOO'
59	'tee'	'ghoo'	'vah'
60	'sah'	'too'	'vee'

61	'see'	'soo'	'ghoo'
62	'GHOO'	'SAH'	'TEE'
63	'tah'	'tee'	'too'
64	'tah'	'ghoo'	'see'
65	'GHOO'	'TEE'	'SAH'
66	'GHEE'	'SOO'	'TEE'
67	'SOO'	'VOO'	'GHEE'
68	'tee'	'vah'	'ghah'
69	'ghoo'	'tee'	'sah'
70	'TEE'	'GHOO'	'VAH'
71	'TEE'	'GHAH'	'VOO'
72	'tah'	'vah'	'soo'
73	'too'	'see'	'ghah'
74	'VOO'	'SOO'	'GHEE'
75	'vee'	'too'	'ghah'
76	'SEE'	'TEE'	'TAH'
77	'SOO'	'TEE'	'GHAH'
78	'TOO'	'SAH'	'VOO'
79	'TAH'	'TEE'	'TOO'
80	'VOO'	'SEE'	'TAH'
81	'TEE'	'VAH'	'GHAH'
82	'tee'	'ghah'	'voo'
83	'GHAH'	'SOO'	'TEE'
84	'ghoo'	'vah'	'see'
85	'SAH'	'GHOO'	'TAH'
86	'GHEE'	'SAH'	'VOO'
87	'TOO'	'SEE'	'GHAH'
88	'ghah'	'see'	'too'
89	'VOO'	'GHEE'	'TAH'
90	'sah'	'ghoo'	'tah'
91	'voo'	'ghee'	'tah'
92	'ghah'	'too'	'vee'
93	'voo'	'tah'	'see'
94	'too'	'sah'	'voo'
95	'GHEE'	'VAH'	'SOO'
96	'ghah'	'too'	'see'
97	'ghah'	'soo'	'tee'
98	'soo'	'voo'	'ghee'
99	'TAH'	'VAH'	'SOO'
100	'too'	'vee'	'ghee'
101	'VEE'	'TAH'	'GHOO'
102	'SOO'	'TEE'	'VAH'
103	'SEE'	'GHAH'	'VAH'
104	'TOO'	'TAH'	'VEE'
105	'tah'	'vee'	'too'
106	'ghee'	'vah'	'soo'
107	'ghah'	'voo'	'ghoo'
108	'ghee'	'voo'	'sah'
109	'GHOO'	'SAH'	'VEE'
110	'GHEE'	'VOO'	'SAH'
111	'vah'	'ghee'	'too'
112	'tee'	'vah'	'soo'
113	'SEE'	'SOO'	'GHOO'
114	'GHEE'	'GHAH'	'SAH'
115	'SAH'	'SEE'	'VAH'
116	'vee'	'tah'	'ghoo'
117	'GHAH'	'VOO'	'GHOO'
118	'TOO'	'VEE'	'GHAH'
119	'voo'	'see'	'tah'
120	'TOO'	'VEE'	'GHEE'

WORD LIST #4

1	'voo'	'vee'	'vah'
2	'tah'	'ghee'	'voo'
3	'tee'	'ghee'	'too'
4	'GHAH'	'TEE'	'SOO'
5	'GHOO'	'SAH'	'TEE'
6	'sah'	'too'	'ghee'
7	'tee'	'ghoo'	'sah'
8	'VOO'	'SAH'	'GHEE'
9	'TOO'	'TAH'	'VEE'
10	'SOO'	'VAH'	'TEE'
11	'VEE'	'GHAH'	'SOO'
12	'voo'	'vah'	'see'
13	'vah'	'too'	'see'
14	'see'	'voo'	'ghah'
15	'ghah'	'soo'	'vee'
16	'TOO'	'SEE'	'GHAH'
17	'ghoo'	'see'	'tah'
18	'GHEE'	'TOO'	'VAH'
19	'see'	'vah'	'ghoo'
20	'voo'	'sah'	'tah'
21	'VOO'	'GHEE'	'SAH'
22	'tah'	'voo'	'soo'
23	'vee'	'ghoo'	'sah'
24	'GHEE'	'VEE'	'VAH'
25	'see'	'ghah'	'voo'
26	'ghee'	'vee'	'vah'
27	'VAH'	'SEE'	'TOO'
28	'GHAH'	'SOO'	'VEE'
29	'see'	'ghoo'	'tah'
30	'soo'	'vah'	'tee'
31	'VAH'	'GHOO'	'SEE'
32	'VAH'	'TOO'	'SEE'
33	'vee'	'ghah'	'soo'
34	'vah'	'ghoo'	'see'
35	'VOO'	'VEE'	'VAH'
36	'VOO'	'SAH'	'TAH'
37	'tah'	'see'	'ghoo'
38	'GHOO'	'SEE'	'TAH'
39	'too'	'see'	'ghah'
40	'VEE'	'SAH'	'GHOO'
41	'GHOO'	'TEE'	'SAH'
42	'vee'	'soo'	'ghah'
43	'soo'	'ghah'	'voo'
44	'SEE'	'VAH'	'TOO'
45	'see'	'vah'	'too'
46	'too'	'tah'	'vah'
47	'SOO'	'TAH'	'VEE'
48	'SEE'	'GHAH'	'VOO'
49	'vee'	'sah'	'ghoo'
50	'vee'	'too'	'tah'
51	'tah'	'ghoo'	'vee'
52	'tee'	'soo'	'ghah'
53	'ghoo'	'tee'	'sah'
54	'sah'	'tee'	'ghoo'
55	'TOO'	'GHEE'	'VOO'
56	'TAH'	'GHEE'	'VOO'
57	'ghee'	'tah'	'tee'
58	'TAH'	'SEE'	'GHOO'
59	'tah'	'soo'	'tee'
60	'ghee'	'too'	'vah'

61	'vah'	'too'	'ghee'
62	'SEE'	'VOO'	'GHAH'
63	'SAH'	'TEE'	'GHOO'
64	'SAH'	'VEE'	'SOO'
65	'GHEE'	'TAH'	'TEE'
66	'TAH'	'VOO'	'SOO'
67	'sah'	'soo'	'tee'
68	'voo'	'sah'	'ghee'
69	'TAH'	'SOO'	'TEE'
70	'GHOO'	'VAH'	'SEE'
71	'TAH'	'GHOO'	'VEE'
72	'ghoo'	'vee'	'sah'
73	'soo'	'ghah'	'tah'
74	'SAH'	'VOO'	'GHEE'
75	'SAH'	'TOO'	'GHEE'
76	'soo'	'tee'	'ghee'
77	'too'	'ghee'	'voo'
78	'GHAH'	'VOO'	'VEE'
79	'soo'	'tah'	'vee'
80	'TEE'	'GHOO'	'SAH'
81	'TOO'	'GHEE'	'VAH'
82	'GHAH'	'VOO'	'SEE'
83	'VEE'	'SOO'	'GHAH'
84	'GHEE'	'SEE'	'GHAH'
85	'ghoo'	'vah'	'see'
86	'ghah'	'tee'	'soo'
87	'SEE'	'GHOO'	'TAH'
88	'GHEE'	'TAH'	'VOO'
89	'TEE'	'SAH'	'GHOO'
90	'too'	'tah'	'vee'
91	'VEE'	'TOO'	'TAH'
92	'GHAH'	'VEE'	'TOO'
93	'vah'	'see'	'too'
94	'tee'	'ghah'	'too'
95	'SOO'	'GHAH'	'TAH'
96	'vah'	'tee'	'soo'
97	'VAH'	'TEE'	'SOO'
98	'ghee'	'see'	'ghah'
99	'ghoo'	'sah'	'tee'
100	'SOO'	'GHAH'	'VOO'
101	'SOO'	'TEE'	'GHEE'
102	'VAH'	'TOO'	'GHEE'
103	'VOO'	'VAH'	'SEE'
104	'voo'	'ghee'	'sah'
105	'tee'	'sah'	'ghoo'
106	'VEE'	'GHOO'	'SAH'
107	'ghah'	'voo'	'vee'
108	'ghah'	'vee'	'too'
109	'GHOO'	'VEE'	'SAH'
110	'sah'	'voo'	'ghee'
111	'TEE'	'GHAH'	'TOO'
112	'SAH'	'SOO'	'TEE'
113	'TEE'	'SOO'	'GHAH'
114	'too'	'ghee'	'vah'
115	'SEE'	'VAH'	'GHOO'
116	'ghah'	'voo'	'see'
117	'TOO'	'TAH'	'VAH'
118	'sah'	'vee'	'soo'
119	'TEE'	'GHEE'	'TOO'
120	'ghee'	'tah'	'voo'

WORD LIST #5

1	'TAH'	'SEE'	'VOO'
2	'ghoo'	'sah'	'vee'
3	'ghah'	'voo'	'tee'
4	'VAH'	'SEE'	'GHOO'
5	'ghah'	'vee'	'soo'
6	'TOO'	'VAH'	'SAH'
7	'VEE'	'GHAH'	'TOO'
8	'TOO'	'VOO'	'TAH'
9	'VOO'	'TAH'	'GHEE'
10	'SEE'	'GHOO'	'VAH'
11	'tee'	'vee'	'sah'
12	'tee'	'soo'	'ghah'
13	'SEE'	'VAH'	'GHOO'
14	'vah'	'soo'	'tee'
15	'GHOO'	'SAH'	'VEE'
16	'voo'	'ghoo'	'tah'
17	'tee'	'voo'	'ghah'
18	'sah'	'tee'	'see'
19	'sah'	'ghoo'	'soo'
20	'vee'	'sah'	'ghoo'
21	'ghee'	'too'	'tah'
22	'GHEE'	'GHAH'	'TAH'
23	'see'	'vah'	'ghoo'
24	'too'	'voo'	'tah'
25	'vah'	'see'	'ghee'
26	'soo'	'ghee'	'vah'
27	'sah'	'tee'	'soo'
28	'GHEE'	'TOO'	'TAH'
29	'tah'	'vee'	'ghoo'
30	'vee'	'too'	'sah'
31	'GHAH'	'VOO'	'TEE'
32	'TEE'	'SOO'	'GHAH'
33	'SOO'	'GHEE'	'VAH'
34	'TEE'	'SEE'	'SOO'
35	'too'	'tee'	'soo'
36	'GHEE'	'TAH'	'VOO'
37	'GHEE'	'SEE'	'VAH'
38	'VAH'	'SOO'	'TEE'
39	'ghah'	'soo'	'vee'
40	'SAH'	'TEE'	'SOO'
41	'VAH'	'SAH'	'TOO'
42	'VOO'	'TAH'	'VAH'
43	'VEE'	'SOO'	'TAH'
44	'SOO'	'GHAH'	'VEE'
45	'voo'	'sah'	'tee'
46	'TAH'	'GHOO'	'TEE'
47	'ghoo'	'sah'	'ghee'
48	'VOO'	'GHOO'	'TAH'
49	'ghee'	'tah'	'voo'
50	'SAH'	'TEE'	'SEE'
51	'SEE'	'TOO'	'GHAH'
52	'SAH'	'GHOO'	'VOO'
53	'VEE'	'TOO'	'SAH'
54	'SAH'	'TEE'	'GHOO'
55	'voo'	'tah'	'ghee'
56	'too'	'vee'	'ghah'
57	'vah'	'sah'	'too'
58	'see'	'ghah'	'too'
59	'SOO'	'TAH'	'GHEE'
60	'GHOO'	'VAH'	'TEE'

61	'TEE'	'VEE'	'SAH'
62	'vah'	'see'	'ghoo'
63	'ghee'	'see'	'vah'
64	'TAH'	'VOO'	'VAH'
65	'sah'	'tee'	'ghoo'
66	'sah'	'ghoo'	'voo'
67	'voo'	'ghah'	'see'
68	'TEE'	'VOO'	'GHAH'
69	'SOO'	'TOO'	'SEE'
70	'ghoo'	'vah'	'see'
71	'vah'	'ghee'	'see'
72	'VEE'	'SAH'	'GHOO'
73	'GHOO'	'VAH'	'SEE'
74	'see'	'ghoo'	'vah'
75	'vee'	'soo'	'tah'
76	'tee'	'see'	'soo'
77	'TEE'	'VAH'	'SAH'
78	'VAH'	'SEE'	'GHEE'
79	'VOO'	'SAH'	'TEE'
80	'GHAH'	'VEE'	'SOO'
81	'vee'	'ghah'	'too'
82	'TOO'	'VEE'	'GHAH'
83	'TAH'	'VEE'	'GHOO'
84	'ghoo'	'tah'	'vee'
85	'ghoo'	'vah'	'tee'
86	'GHAH'	'GHEE'	'VOO'
87	'tah'	'see'	'voo'
88	'too'	'tee'	'ghee'
89	'GHEE'	'TOO'	'SAH'
90	'see'	'voo'	'ghah'
91	'TAH'	'SOO'	'VEE'
92	'tah'	'soo'	'vee'
93	'SOO'	'GHEE'	'VOO'
94	'SEE'	'VOO'	'GHAH'
95	'ghah'	'ghee'	'voo'
96	'GHAH'	'SOO'	'VEE'
97	'GHAH'	'VEE'	'TOO'
98	'VEE'	'GHEE'	'TOO'
99	'TOO'	'TEE'	'GHEE'
100	'GHOO'	'TAH'	'VEE'
101	'soo'	'ghah'	'vee'
102	'vee'	'ghee'	'too'
103	'too'	'vah'	'sah'
104	'tah'	'voo'	'vah'
105	'VAH'	'GHEE'	'SEE'
106	'SEE'	'GHAH'	'TOO'
107	'soo'	'tah'	'ghee'
108	'VOO'	'GHAH'	'SEE'
109	'SAH'	'GHOO'	'SOO'
110	'ghah'	'vee'	'too'
111	'see'	'too'	'ghah'
112	'ghee'	'too'	'sah'
113	'tah'	'ghoo'	'tee'
114	'ghee'	'ghah'	'tah'
115	'GHOO'	'SAH'	'GHEE'
116	'soo'	'ghee'	'voo'
117	'soo'	'too'	'see'
118	'tee'	'vah'	'sah'
119	'TOO'	'TEE'	'SOO'
120	'voo'	'tah'	'vah'

Appendix B Supplemental Information and Figures

This appendix includes more detailed information about data analysis. Figures and descriptions that were not appropriate for the proposed manuscript sections are included below. Supplemental information primarily relates to manuscript #1, focusing on phonetic-intensity encoding or the analysis of articulatory precision.

Vocal Intensity

Prior to analysis, data were screened for acoustic contamination by referencing notations of background noise and non-speech patient noise (i.e. coughing) made during coding. Vocal intensity values (normalized RMS) which were above the 75^{th} percentile of the distribution and contained a notation of background noise or non-patient speech were excluded from analysis. Less than 1% of the data were excluded based on acoustic contamination. The distribution was examined to determine if the participants executed the loud/soft variation in distinct groups (i.e. bimodal distribution) or a continuous distribution (Figure 20). The data were normally distributed based on visualization and did not display a bimodal distribution, indicating that the participants produced a range of intensities, but not two distinct groups. This pattern was consistent across participants. Because participants did not produce two distinct intensity levels, correlation was determined to be most appropriate analysis method to assess changes associated with increases in vocal intensity.

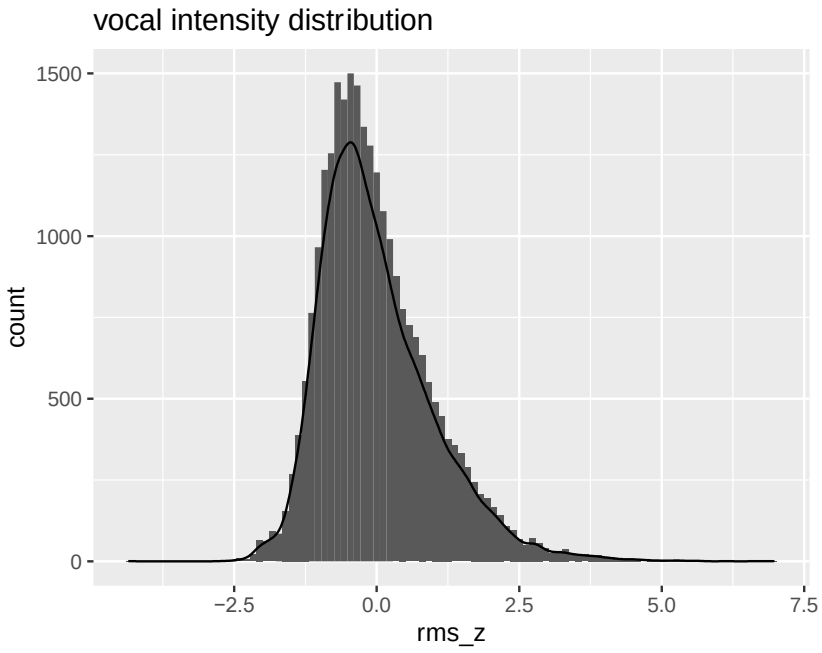


Figure 20. Histogram of Vocal Intensity Values Across All Participants

Temporal Measures of Articulation

Fricative duration (/s/ and /v/) and vowel duration were correlated to vocal intensity to evaluate temporal articulatory changes which occurred with increasing vocal intensity. Significant Pearson correlations were observed for all three measurements for the combined group statistic: /s/ duration ($r = -0.073$, $p < 0.001$), /v/ duration ($r = -0.16$, $p < 0.001$), vowel duration ($r = 0.11$, $p < 0.001$). Outliers were observed in the /s/ and /v/ duration measures, with two participants (DBS3014, DBS3025) producing multiple durations greater than 500 ms. These group correlations were computed excluding the participants with outliers, and the measurements remained significant for both fricative correlations. Similarly, three participants (DBS3004, DBS3015, DBS3006) demonstrated vowel durations longer than 500 ms; therefore, the statistic was computed excluding these participants and it also remained significant. Individual correlations for each participant were computed to determine if the group pattern observed was consistent within individuals. In

all three measurements, the majority of the effect estimates match the direction of the group estimate, with large portions of the individual correlations displaying significant results after Bonferroni correction (Figure 11 B,D,F). Although significant trend lines were identified in all measurements, the scatterplot demonstrates a wide range of variability in productions. To evaluate this variability further, behavioral data were qualitatively examined by syllable position to explain more of the variance. However, for all measures, the spread of the data was nearly identical for all syllable positions. To determine if a specific vowel was driving the results for the correlation with vowel duration, the vowels were examined separately. Four out of the five vowels examined demonstrated significant positive correlations with vocal intensity, mirroring the combined group statistic (Figure 21).

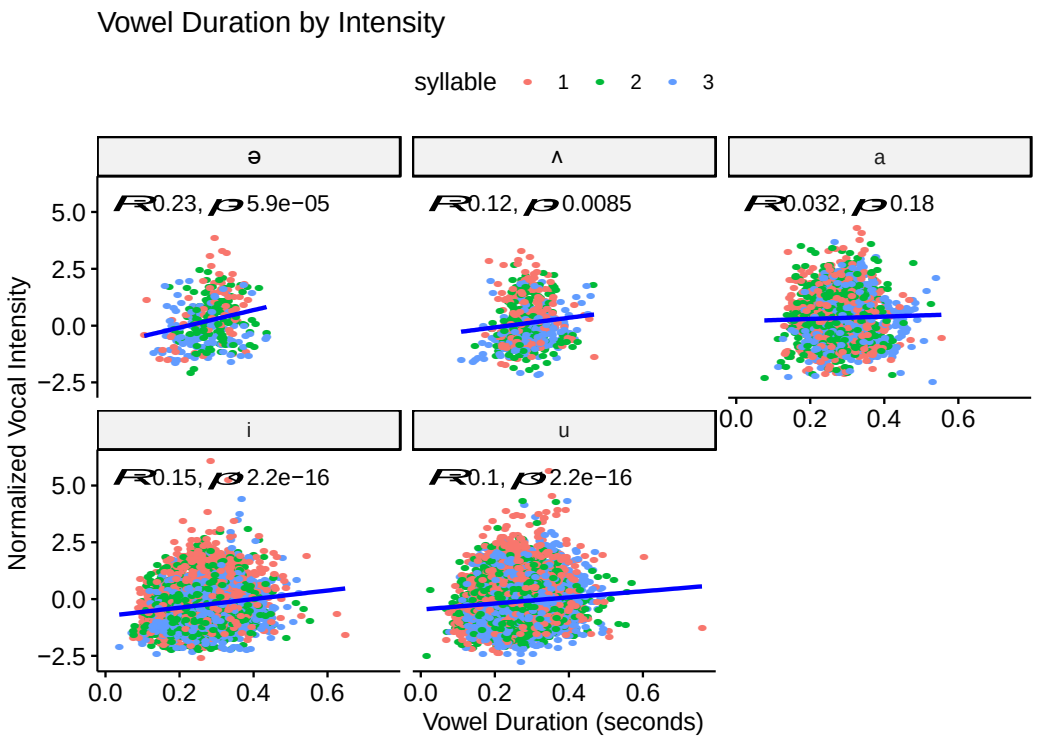


Figure 21. Vowel duration correlates to vocal intensity in multiple vowels.

Spectral Measures

Two spectral measures, the second formant ratio (F2 Ratio) and the spectral centroid frequency, were correlated to vocal intensity to evaluate spectral articulatory changes which occurred with increasing vocal intensity. Group correlation demonstrated a significant effect between vocal intensity and spectral centroid for /s/ ($r = 0.1$, $p < 0.001$), but not for /t/ ($r = 0.017$, $p = 0.15$). When inspecting results stratified by participant for the /s/ phoneme, high spectral centroid outliers appeared to be driving the correlation from a single participant (DBS3032) (Figure 22). The correlation was repeated after excluding this participant, and the behavioral result remained significant and consistent with the overall result ($r = 0.19$, $p < 0.001$).

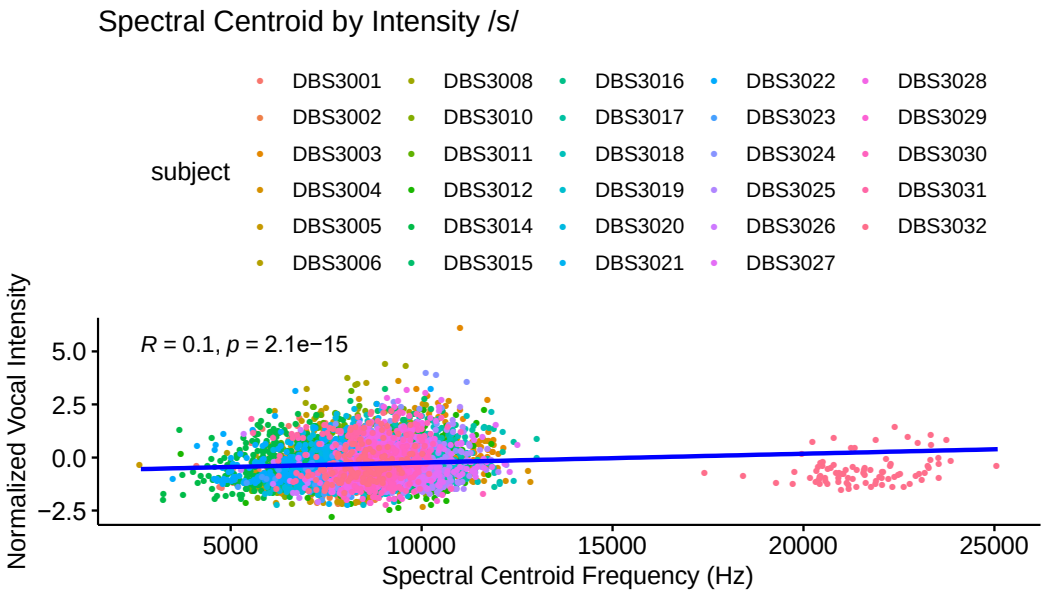


Figure 22. Spectral centroid frequency outliers.

A similar outlier was observed for spectral centroid of the /t/ productions; therefore, DBS3032 was excluded from the correlation. This exclusion caused the overall finding to become significant for /t/ ($r = 0.035$, $p = 0.0041$). Individual correlations reflected the group results. Most participants

demonstrated a significant positive relationship between vocal intensity and spectral centroid frequency for the /s/ phoneme after Bonferroni correction (69%, 20/29), but only 14% (6/29) demonstrated a significant relationship for the /t/ phoneme.

Normality Assessment of Subthalamic Nucleus Theta Power

Neural power in STN theta band, which is the primary effect of interest in this model were fit with a normal distribution assumption for the mixed-effects model; however, this assumption was not strictly met. Figure 23 below demonstrates data skewedness. A second model was performed to test that those effects were not identified due to randomness in a large sample size. The model was refit using randomized labels for electrode location, which resulted in no significant effects, confirming that effects were not a product of random behavior.

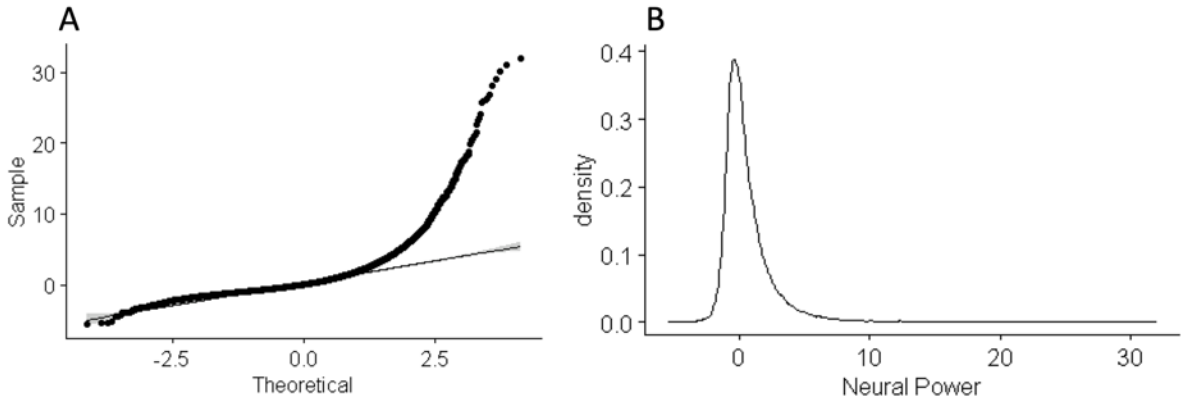


Figure 23 Subthalamic Nucleus (STN) Theta Power Normality Assessment

A. QQ plot of the STN neural power with theoretical normally distributed data plotted on the x axis and actual data plotted on the y axis. The non-linear behavior in the positive range indicates a positive skew. **B.** Density plot of STN neural power, demonstrating a long positive tail in the distribution.

Articulator Labeling Performance Assessment

Articulator labeling was designed to assign an articulator label to an electrode if power was significantly greater when that articulator was in use compared to other electrodes. The figure below depicts the significant articulator by phoneme interaction which was identified in the final mixed-effects model evaluating the three-way interaction between phoneme, articulatory and intensity. These results verify that the articulator labels performed as intended.

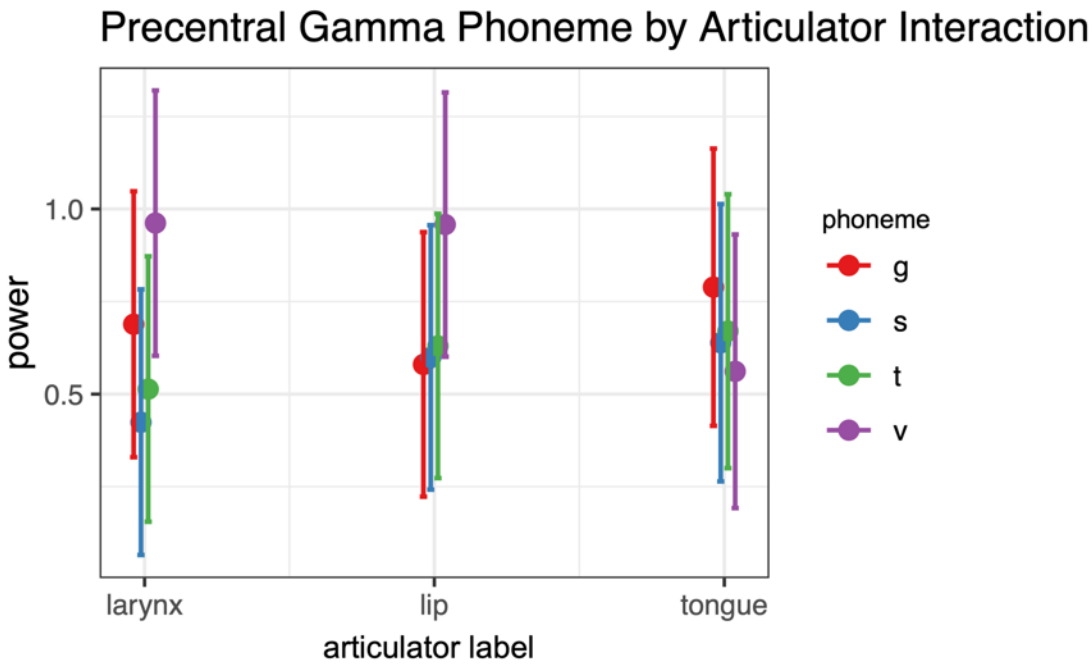
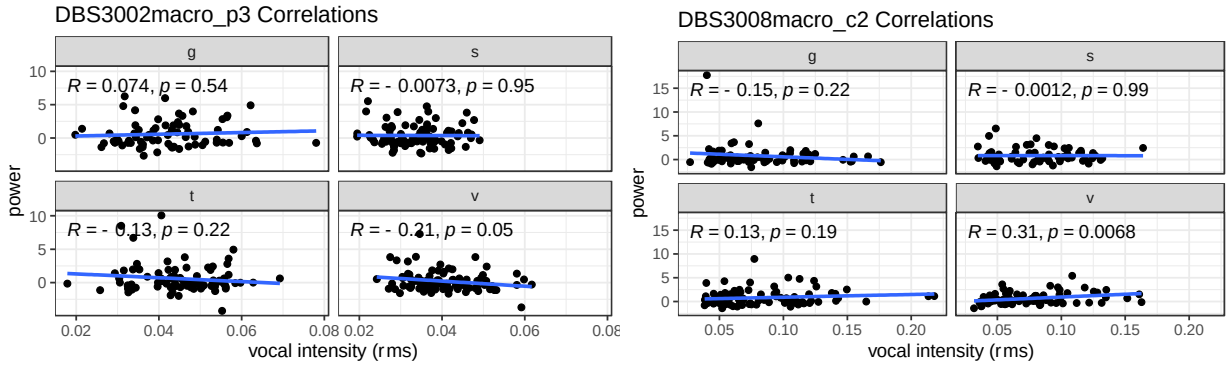


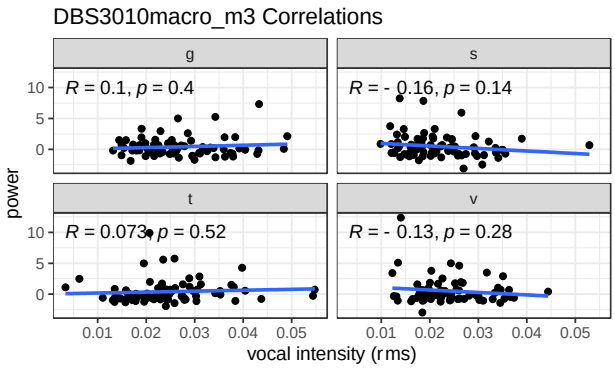
Figure 24 Articulator Labeling Model Performance

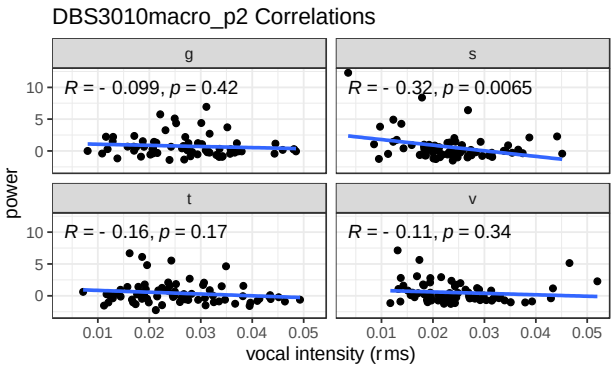
A mixed-effects model was fit evaluating the three-way interaction between phoneme, articulatory and intensity, which also produced an articulator by phoneme interaction effect. Effect estimates reflect the intended performance of articulator labels. For example, the lip articulator is only used in the /v/ phoneme, and the power effect estimate is higher compared to the remaining phonemes which do not use the lip articulator.

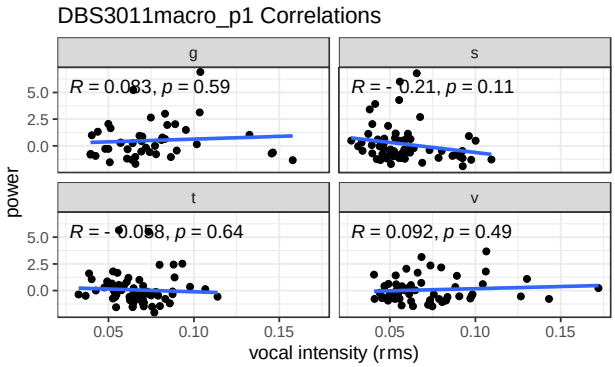
Correlation of Vocal Intensity to STN Theta Band Power

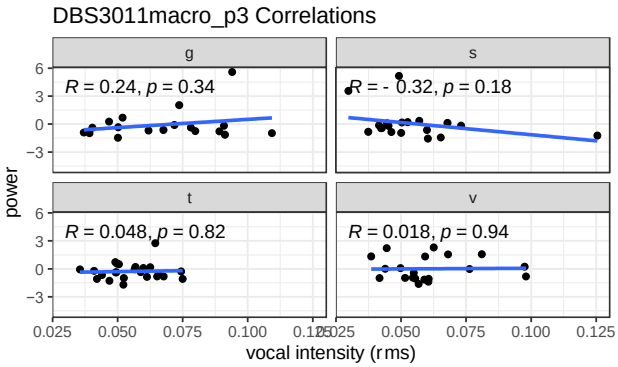
A significant three-way interaction effect between power, phoneme and recording location was identified in the STN theta band power, indicating the presence of *phonetic-intensity encoding*. A total of 24 electrodes displayed at least one significant contrast in the model summary; however, data from many recording locations would not converge during post-hoc simple slope analysis. To further understand patterns within these significant recording locations, Pearson correlations were conducted between vocal intensity and STN theta band power, stratified by recording location and phoneme. Correlations are included below.

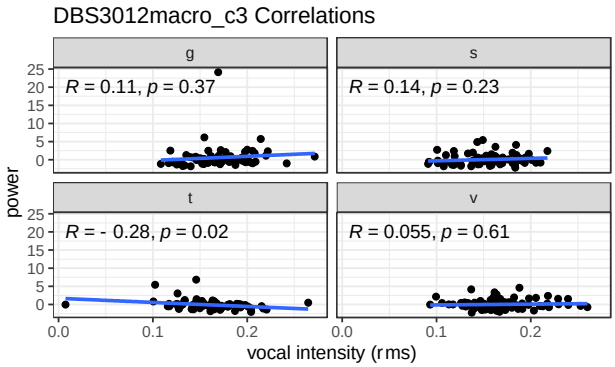


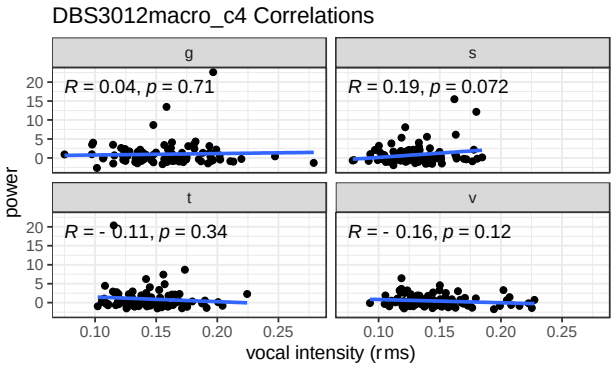

DBS3010macro_m3 Correlations
g
s
t
v
R = 0.1, p = 0.4
R = - 0.16, p = 0.14
R = 0.073, p = 0.52
R = - 0.13, p = 0.28
power
vocal intensity (rms)

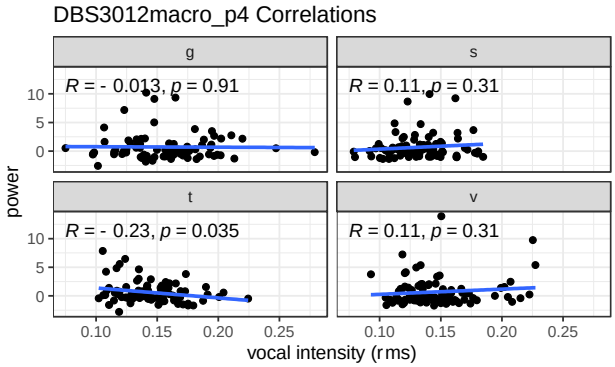

DBS3010macro_p2 Correlations
g
s
t
v
R = - 0.099, p = 0.42
R = - 0.32, p = 0.0065
R = - 0.16, p = 0.17
R = - 0.11, p = 0.34
power
vocal intensity (rms)

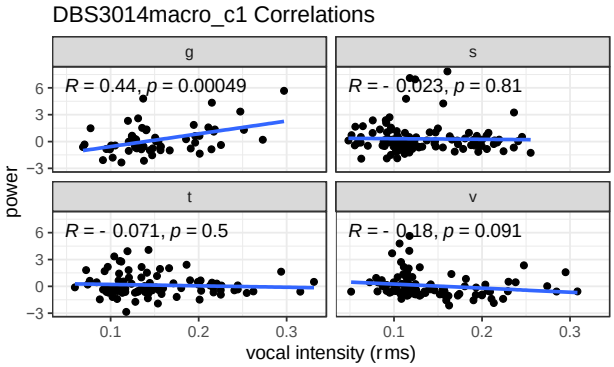

DBS3011macro_p1 Correlations
g
s
t
v
R = 0.083, p = 0.59
R = - 0.21, p = 0.11
R = - 0.058, p = 0.64
R = 0.092, p = 0.49
power
vocal intensity (rms)

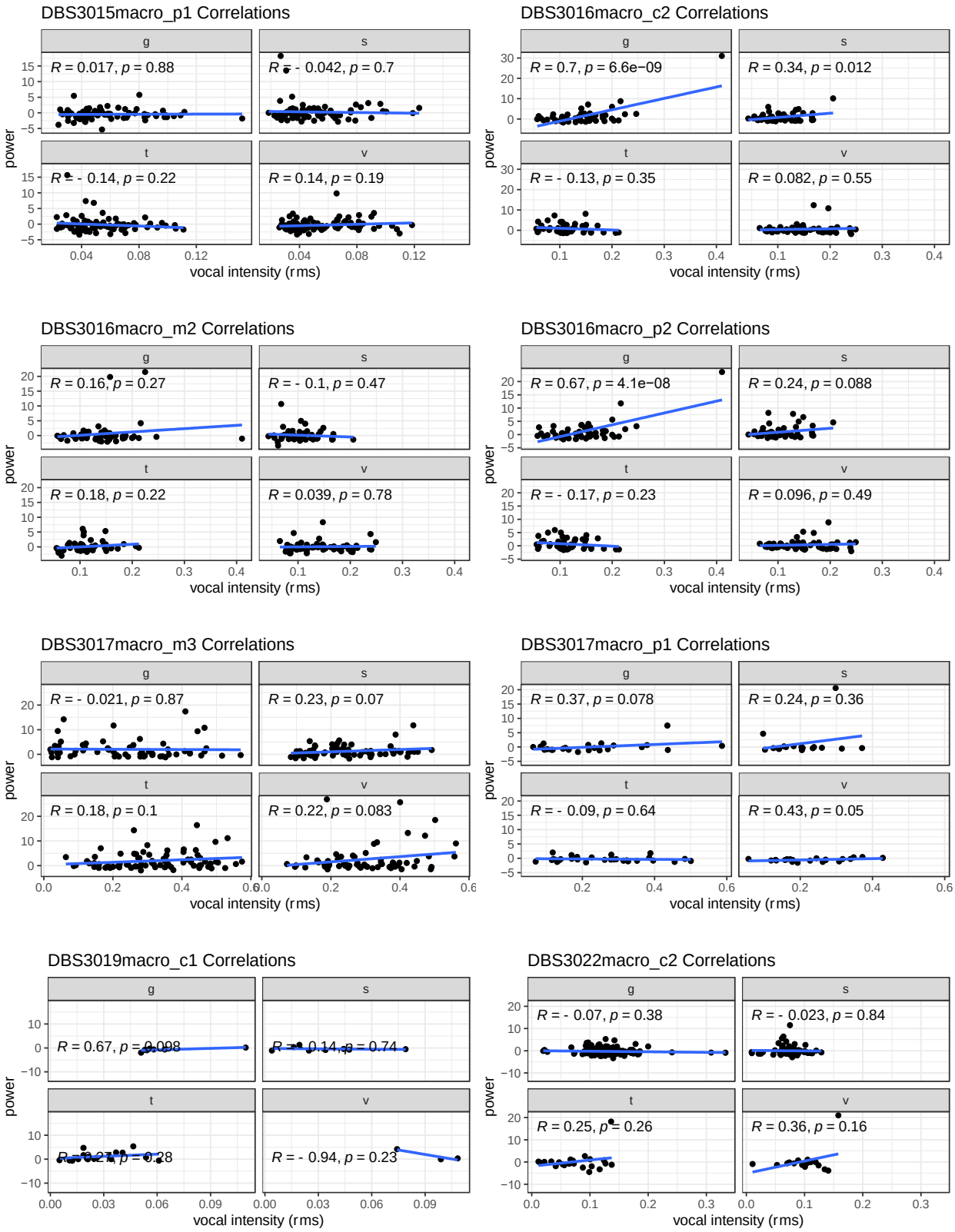

DBS3011macro_p3 Correlations
g
s
t
v
R = 0.24, p = 0.34
R = - 0.32, p = 0.18
R = 0.048, p = 0.82
R = 0.018, p = 0.94
power
vocal intensity (rms)

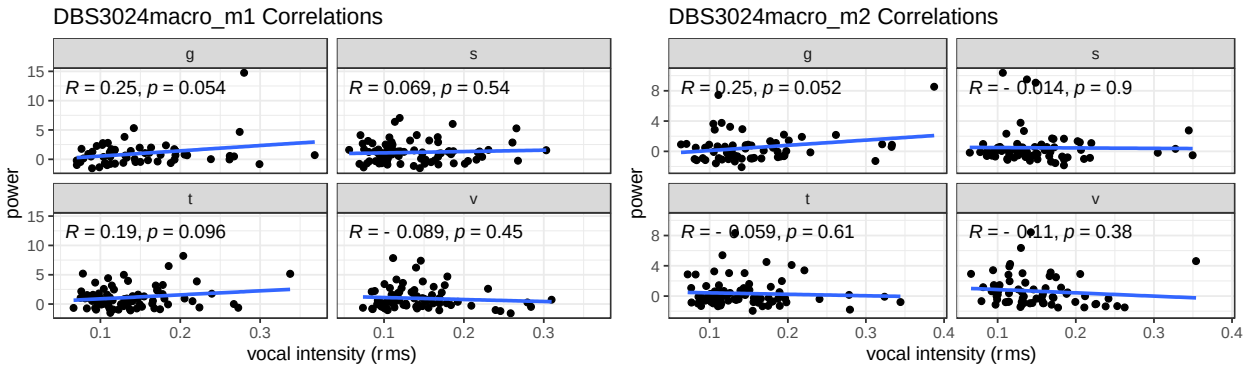

DBS3012macro_c3 Correlations
g
s
t
v
R = 0.11, p = 0.37
R = 0.14, p = 0.23
R = - 0.28, p = 0.02
R = 0.055, p = 0.61
power
vocal intensity (rms)


DBS3012macro_c4 Correlations
g
s
t
v
R = 0.04, p = 0.71
R = 0.19, p = 0.072
R = - 0.11, p = 0.34
R = - 0.16, p = 0.12
power
vocal intensity (rms)

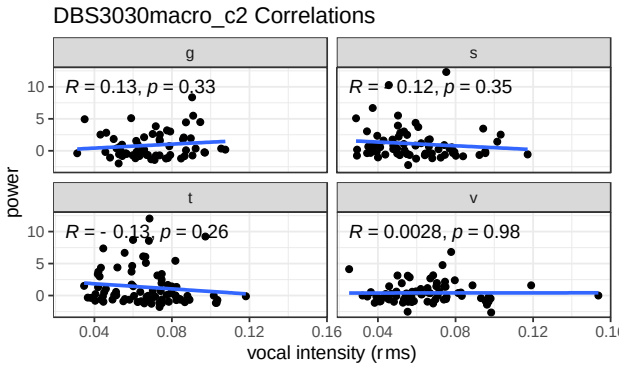

DBS3012macro_p4 Correlations
g
s
t
v
R = - 0.013, p = 0.91
R = 0.11, p = 0.31
R = - 0.23, p = 0.035
R = 0.11, p = 0.31
power
vocal intensity (rms)

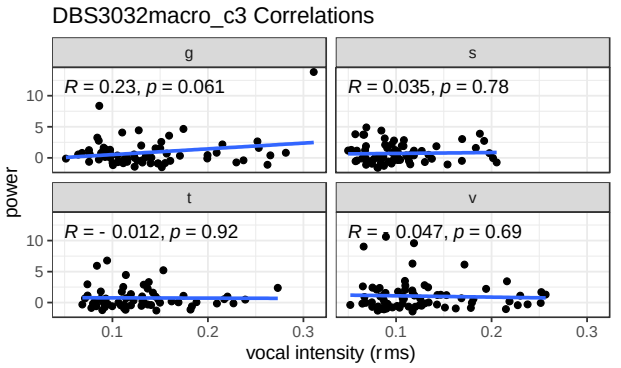

DBS3014macro_c1 Correlations
g
s
t
v
R = 0.44, p = 0.00049
R = - 0.023, p = 0.81
R = - 0.071, p = 0.5
R = - 0.18, p = 0.091
power
vocal intensity (rms)

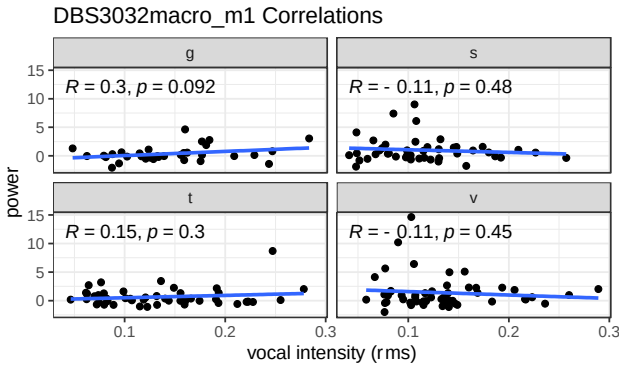

DBS3015macro_p1 Correlations
g
s
t
v
R = 0.017, p = 0.88
R = - 0.042, p = 0.7
R = - 0.14, p = 0.22
R = 0.14, p = 0.19
power
vocal intensity (rms)
DBS3016macro_c2 Correlations
R = 0.7, p = 6.6e−09
R = 0.34, p = 0.012
R = - 0.13, p = 0.35
R = 0.082, p = 0.55
DBS3016macro_m2 Correlations
R = 0.16, p = 0.27
R = - 0.1, p = 0.47
R = 0.18, p = 0.22
R = 0.039, p = 0.78
DBS3016macro_p2 Correlations
R = 0.67, p = 4.1e−08
R = 0.24, p = 0.088
R = - 0.17, p = 0.23
R = 0.096, p = 0.49
DBS3017macro_m3 Correlations
R = - 0.021, p = 0.87
R = 0.23, p = 0.07
R = 0.18, p = 0.1
R = 0.22, p = 0.083
DBS3017macro_p1 Correlations
R = 0.37, p = 0.078
R = 0.24, p = 0.36
R = - 0.09, p = 0.64
R = 0.43, p = 0.05
DBS3019macro_c1 Correlations
R = 0.67, p = 0.098
R = - 0.14, p = 0.74
R = 0.27, p = 0.28
R = - 0.94, p = 0.23
DBS3022macro_c2 Correlations
R = - 0.07, p = 0.38
R = - 0.023, p = 0.84
R = 0.25, p = 0.26
R = 0.36, p = 0.16

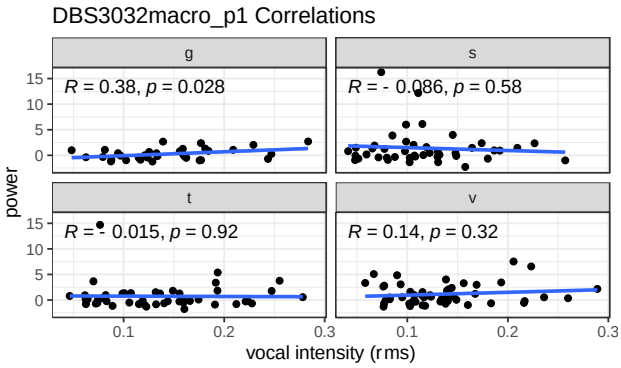

DBS3024macro_m1 Correlations
g
s
t
v
R = 0.25, p = 0.054
R = 0.069, p = 0.54
R = 0.19, p = 0.096
R = - 0.089, p = 0.45
power
vocal intensity (rms)

DBS3024macro_m2 Correlations
g
s
t
v
R = 0.25, p = 0.052
R = - 0.014, p = 0.9
R = - 0.059, p = 0.61
R = - 0.11, p = 0.38
power
vocal intensity (rms)


DBS3030macro_c2 Correlations
g
s
t
v
R = 0.13, p = 0.33
R = - 0.12, p = 0.35
R = - 0.13, p = 0.26
R = 0.0028, p = 0.98
power
vocal intensity (rms)


DBS3032macro_c3 Correlations
g
s
t
v
R = 0.23, p = 0.061
R = 0.035, p = 0.78
R = - 0.012, p = 0.92
R = 0.047, p = 0.69
power
vocal intensity (rms)


DBS3032macro_m1 Correlations
g
s
t
v
R = 0.3, p = 0.092
R = - 0.11, p = 0.48
R = 0.15, p = 0.3
R = - 0.11, p = 0.45
power
vocal intensity (rms)


DBS3032macro_p1 Correlations
g
s
t
v
R = 0.38, p = 0.028
R = - 0.086, p = 0.58
R = - 0.015, p = 0.92
R = 0.14, p = 0.32
power
vocal intensity (rms)

Bibliography

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends Neurosci, 12*(10), 366-375. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2479133

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 Relat Disord, 33*, 3-11. doi:10.1016/j.parkreldis.2016.09.022

Alegre, M., Lopez-Azcarate, J., Obeso, I., Wilkinson, L., Rodriguez-Oroz, M. C., Valencia, M., . . . Obeso, J. A. (2013). The subthalamic nucleus is involved in successful inhibition in the stop-signal task: a local field potential study in Parkinson's disease. *Exp Neurol, 239*, 1-12. doi:10.1016/j.expneurol.2012.08.027

Alexander, G. E., Crutcher, M. D., & DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog Brain Res, 85*, 119-146. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2094891

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci, 9*, 357-381. doi:10.1146/annurev.ne.09.030186.002041

Androulidakis, A. G., Doyle, L. M., Yarrow, K., Litvak, V., Gilbertson, T. P., & Brown, P. (2007). Anticipatory changes in beta synchrony in the human corticospinal system and associated improvements in task performance. *Eur J Neurosci, 25*(12), 3758-3765. doi:10.1111/j.1460-9568.2007.05620.x

Anzak, A., Gaynor, L., Beigi, M., Foltynie, T., Limousin, P., Zrinzo, L., . . . Jahanshahi, M. (2013). Subthalamic nucleus gamma oscillations mediate a switch from automatic to controlled processing: a study of random number generation in Parkinson's disease. *Neuroimage, 64*, 284-289. doi:10.1016/j.neuroimage.2012.08.068

Anzak, A., Tan, H., Pogosyan, A., Foltynie, T., Limousin, P., Zrinzo, L., . . . Brown, P. (2012). Subthalamic nucleus activity optimizes maximal effort motor responses in Parkinson's disease. *Brain, 135*(Pt 9), 2766-2778. doi:10.1093/brain/aws183

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends Cogn Sci, 16*(7), 390-398. doi:10.1016/j.tics.2012.05.003

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *J Neurosci, 26*(9), 2424-2433. doi:10.1523/JNEUROSCI.4682-05.2006

Astrom, M., Tripoliti, E., Hariz, M. I., Zrinzo, L. U., Martinez-Torres, I., Limousin, P., & Wardell, K. (2010). Patient-specific model-based investigation of speech intelligibility and movement during deep brain stimulation. *Stereotact Funct Neurosurg, 88*(4), 224-233. doi:10.1159/000314357

Aziz-Zadeh, L., Sheng, T., & Gheytanchi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS One, 5*(1), e8759. doi:10.1371/journal.pone.0008759

Baker, K. K., Ramig, L. O., Sapir, S., Luschei, E. S., & Smith, M. E. (2001). Control of vocal loudness in young and old adults. *J Speech Lang Hear Res, 44*(2), 297-305. doi:10.1044/1092-4388(2001/024)

Bakhtiari, R., Cummine, J., Reed, A., Fox, C. M., Chouinard, B., Cribben, I., & Boliek, C. A. (2017). Changes in brain activity following intensive voice treatment in children with cerebral palsy. *Hum Brain Mapp, 38*(9), 4413-4429. doi:10.1002/hbm.23669

Ball, T., Demandt, E., Mutschler, I., Neitzel, E., Mehring, C., Vogt, K., . . . Schulze-Bonhage, A. (2008). Movement related activity in the high gamma range of the human EEG. *Neuroimage, 41*(2), 302-310. doi:10.1016/j.neuroimage.2008.02.032

Ballard, K. J., Robin, D. A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *J Speech Lang Hear Res, 53*(5), 1227-1245. doi:10.1044/1092-4388(2010/09-0130)

Bar-Gad, I., Morris, G., & Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog Neurobiol, 71*(6), 439-473. doi:10.1016/j.pneurobio.2003.12.001

Barbosa, P. (2007). From syntax to acoustic duration: A dynamical model of speech rhythm production. *Speech Communication, 49*, 725-742.

Barnish, M. S., Horton, S. M. C., Butterfint, Z. R., Clark, A. B., Atkinson, R. A., & Deane, K. H. O. (2017). Speech and communication in Parkinson's disease: a cross-sectional exploratory study in the UK. *BMJ Open, 7*(5), e014642. doi:10.1136/bmjopen-2016-014642

Basinger, H., & Joseph, J. (2021). Neuroanatomy, Subthalamic Nucleus. In *StatPearls*. Treasure Island (FL).

Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody: Insights from lesions studies and neuroimaging. . *Aphasiology, 13*(8), 581-608.

Baumann, A., Nebel, A., Granert, O., Giehl, K., Wolff, S., Schmidt, W., . . . van Eimeren, T. (2018). Neural Correlates of Hypokinetic Dysarthria and Mechanisms of Effective Voice Treatment in Parkinson Disease. *Neurorehabil Neural Repair, 32*(12), 1055-1066. doi:10.1177/1545968318812726

Baunez, C., Yelnik, J., & Mallet, L. (2011). Six questions on the subthalamic nucleus: lessons from animal models and from stimulated patients. *Neuroscience, 198*, 193-204. doi:10.1016/j.neuroscience.2011.09.059

Behrman, A., Cody, J., Elandary, S., Flom, P., & Chitnis, S. (2020). The Effect of SPEAK OUT! and The LOUD Crowd on Dysarthria Due to Parkinson's Disease. *Am J Speech Lang Pathol, 29*(3), 1448-1465. doi:10.1044/2020_AJSLP-19-00024

Behroozmand, R., Johari, K., Kelley, R. M., Kapnoula, E. C., Narayanan, N. S., & Greenlee, J. D. W. (2019). Effect of deep brain stimulation on vocal motor control mechanisms in Parkinson's disease. *Parkinsonism Relat Disord*. doi:10.1016/j.parkreldis.2019.03.002

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *J Cogn Neurosci, 22*(7), 1504-1529. doi:10.1162/jocn.2009.21306

Bostan, A. C., Dum, R. P., & Strick, P. L. (2010). The basal ganglia communicate with the cerebellum. *Proc Natl Acad Sci U S A, 107*(18), 8452-8456. doi:10.1073/pnas.1000496107

Bouchard, K. E., & Chang, E. F. (2014a). Control of spoken vowel acoustics and the influence of phonetic context in human speech sensorimotor cortex. *J Neurosci, 34*(38), 12662-12677. doi:10.1523/JNEUROSCI.1219-14.2014

Bouchard, K. E., & Chang, E. F. (2014b). Neural decoding of spoken vowels from human sensory-motor cortex with high-density electrocorticography. *Conf Proc IEEE Eng Med Biol Soc, 2014*, 6782-6785. doi:10.1109/EMBC.2014.6945185

Bouchard, K. E., Conant, D. F., Anumanchipalli, G. K., Dichter, B., Chaisanguanthum, K. S., Johnson, K., & Chang, E. F. (2016). High-Resolution, Non-Invasive Imaging of Upper Vocal Tract Articulators Compatible with Human Brain Recordings. *PLoS One, 11*(3), e0151327. doi:10.1371/journal.pone.0151327

Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature, 495*(7441), 327-332. doi:10.1038/nature11911

Boutsen, F., Park, E., Dvorak, J., & Cid, C. (2018). Prosodic Improvement in Persons with Parkinson Disease Receiving SPEAK OUT!(R) Voice Therapy. *Folia Phoniatr Logop, 70*(2), 51-58. doi:10.1159/000488875

Brittain, J. S., & Brown, P. (2014). Oscillations and the basal ganglia: motor control and beyond. *Neuroimage, 85 Pt 2*, 637-647. doi:10.1016/j.neuroimage.2013.05.084

Browman, C., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*, 151-206.

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Hum Brain Mapp, 25*(1), 105-117. doi:10.1002/hbm.20140

Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2009). The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain Cogn, 70*(1), 31-41. doi:10.1016/j.bandc.2008.12.006

Brucke, C., Huebl, J., Schonecker, T., Neumann, W. J., Yarrow, K., Kupsch, A., . . . Kuhn, A. A. (2012). Scaling of movement is related to pallidal gamma oscillations in patients with dystonia. *J Neurosci, 32*(3), 1008-1019. doi:10.1523/JNEUROSCI.3860-11.2012

Brucke, C., Kempf, F., Kupsch, A., Schneider, G. H., Krauss, J. K., Aziz, T., . . . Kuhn, A. A. (2008). Movement-related synchronization of gamma activity is lateralized in patients with dystonia. *Eur J Neurosci, 27*(9), 2322-2329. doi:10.1111/j.1460-9568.2008.06203.x

Bush, A. (2021). brain-modulation-lab/bml. *GitHub Repository.* Retrieved from https://github.com/Brain-Modulation-Lab/bml

Bush, A., Chrabaszcz, A., Peterson, V., Dastolfo-Hromack, C., Saravanan, V., & Lipski, W. (2021). *Identification and mitigation of technical challenges in experiments probing the cortical-basal ganglia speech network during awake Deep Brain Stimulation surgeries.* Paper presented at the Society for Neuroscience Global Connectome, Virtual.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics, 31*, 149-180.

Caekebeke, J. F., Jennekens-Schinkel, A., van der Linden, M. E., Buruma, O. J., & Roos, R. A. (1991). The interpretation of dysprosody in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry, 54*(2), 145-148. doi:10.1136/jnnp.54.2.145

Cancelliere, A. E., & Kertesz, A. (1990). Lesion localization in acquired deficits of emotional expression and comprehension. *Brain Cogn, 13*(2), 133-147. doi:10.1016/0278-2626(90)90046-q

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., . . . Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science, 313*(5793), 1626-1628. doi:10.1126/science.1128115

Canter, G. J. (1965). Speech Characteristics of Patients with Parkinson's Disease. 3. Articulation, Diadochokinesis, and over-All Speech Adequacy. *J Speech Hear Disord, 30*, 217-224. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14321495

Carmichael, C., Sapienza, C., & Ramig, L. O. (2006). *EMST and LSVT in Patients with Parkinson's Disease.* Paper presented at the Speech Motor Control Conference, Austin, TX.

Carpenter, M. B., Baton, R. R., 3rd, Carleton, S. C., & Keller, J. T. (1981). Interconnections and organization of pallidal and subthalamic nucleus neurons in the monkey. *J Comp Neurol, 197*(4), 579-603. doi:10.1002/cne.901970404

Cassidy, M., Mazzone, P., Oliviero, A., Insola, A., Tonali, P., Di Lazzaro, V., & Brown, P. (2002). Movement-related changes in synchronization in the human basal ganglia. *Brain, 125*(Pt 6), 1235-1246. doi:10.1093/brain/awf135

Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueroa, C. M., Samanta, J., Sherman, S. J., & Frank, M. J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat Neurosci, 14*(11), 1462-1467. doi:10.1038/nn.2925

Cebolla, A. M., & Cheron, G. (2019). Understanding Neural Oscillations in the Human Brain: From Movement to Consciousness and Vice Versa. *Front Psychol, 10*, 1930. doi:10.3389/fpsyg.2019.01930

Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc Natl Acad Sci U S A, 110*(7), 2653-2658. doi:10.1073/pnas.1216827110

Chartier, J., Anumanchipalli, G. K., Johnson, K., & Chang, E. F. (2018). Encoding of Articulatory Kinematic Trajectories in Human Speech Sensorimotor Cortex. *Neuron, 98*(5), 1042-1054 e1044. doi:10.1016/j.neuron.2018.04.031

Chen, Y. C., Thaler, D., Nixon, P. D., Stern, C. E., & Passingham, R. E. (1995). The functions of the medial premotor cortex. II. The timing and selection of learned movements. *Exp Brain Res, 102*(3), 461-473. doi:10.1007/bf00230650

Cheyne, D., Bells, S., Ferrari, P., Gaetz, W., & Bostan, A. C. (2008). Self-paced movements induce high-frequency gamma oscillations in primary motor cortex. *Neuroimage, 42*(1), 332-342. doi:10.1016/j.neuroimage.2008.04.178

Cho, T., & Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*, 155-190.

Chrabaszcz, A., Neumann, W. J., Stretcu, O., Lipski, W. J., Bush, A., Dastolfo-Hromack, C., . . . Richardson, R. M. (2019). Subthalamic nucleus and sensorimotor cortex activity during speech production. *J Neurosci*. doi:10.1523/JNEUROSCI.2842-18.2019

Coburn, E. (2016). *Nonword Repetition Task to Evaluate Syllable Stress as a Motor Class*. (Master of Science). West Virginia University,

Collaborators, G. B. D. P. s. D. (2018). Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol, 17*(11), 939-953. doi:10.1016/S1474-4422(18)30295-3

Conant, D., Bouchard, K. E., & Chang, E. F. (2014). Speech map in the human ventral sensory-motor cortex. *Curr Opin Neurobiol, 24*(1), 63-67. doi:10.1016/j.conb.2013.08.015

Conant, D. F., Bouchard, K. E., Leonard, M. K., & Chang, E. F. (2018). Human Sensorimotor Cortex Control of Directly Measured Vocal Tract Movements during Vowel Production. *J Neurosci, 38*(12), 2955-2966. doi:10.1523/JNEUROSCI.2382-17.2018

Crone, N. E., Miglioretti, D. L., Gordon, B., Sieracki, J. M., Wilson, M. T., Uematsu, S., & Lesser, R. P. (1998). Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. I. Alpha and beta event-related desynchronization. *Brain, 121 (Pt 12)*, 2271-2299. doi:10.1093/brain/121.12.2271

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. *J Voice, 22*(3), 365-372. doi:10.1016/j.jvoice.2006.10.010

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9*(2), 179-194. doi:10.1006/nimg.1998.0395

Damron, R. (2004). Prosodic Schemas. In M. C. & M.-Z. A. (Eds.), *Discourse Across Languages and Cultures*: John Benjamins Publishing Company.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia: Saunders.

Dastolfo-Hromack, C., Alhourani, A., Crammond, D. J., Shaiman, S., Dickey, M., Holt, L., . . . Richardson, R. M. (2018). *Local field potentials in the human subthalamic nucleus predict scaling of speech production*. Paper presented at the Society for Neuroscience, San Diego, CA.

Dastolfo-Hromack, C., Bush, A., Chrabaszcz, A., Alhourani, A., Lipski, W., Wang, D., . . . Richardson, R. M. (2021). Articulatory Gain Predicts Motor Cortex and Subthalamic Nucleus Activity During Speech. *Cerebral Cortex, 00*, 1-13. doi:https://doi.org/10.1093/cercor/bhab251

Degos, B., Deniau, J. M., Le Cam, J., Mailly, P., & Maurice, N. (2008). Evidence for a direct subthalamo-cortical loop circuit in the rat. *Eur J Neurosci, 27*(10), 2599-2610. doi:10.1111/j.1460-9568.2008.06229.x

DeLong, M., & Wichmann, T. (2009). Update on models of basal ganglia function and dysfunction. *Parkinsonism Relat Disord, 15 Suppl 3*, S237-240. doi:10.1016/S1353-8020(09)70822-3

DeLong, M., & Wichmann, T. (2010). Changing views of basal ganglia circuits and circuit disorders. *Clin EEG Neurosci, 41*(2), 61-67. doi:10.1177/155005941004100204

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends Neurosci, 13*(7), 281-285. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1695404

DeLong, M. R., Crutcher, M. D., & Georgopoulos, A. P. (1983). Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *J Neurosci, 3*(8), 1599-1606. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6875659

Dias, A. E., Barbosa, E. R., Coracini, K., Maia, F., Marcolin, M. A., & Fregni, F. (2006). Effects of repetitive transcranial magnetic stimulation on voice and speech in Parkinson's disease. *Acta Neurol Scand, 113*(2), 92-99. doi:10.1111/j.1600-0404.2005.00558.x

Doesburg, S. M., Vinette, S. A., Cheung, M. J., & Pang, E. W. (2012). Theta-modulated gamma-band synchronization among activated regions during a verb generation task. *Front Psychol, 3*, 195. doi:10.3389/fpsyg.2012.00195

Dromey, C. (2010). Laryngeal articulatory coupling in three speech disorders. In B. Maassen & P. Van Lieshout (Eds.), *Speech Motor Control: New developments in basic and applied research*. Oxford; New York: Oxford University Press.

Dromey, C., Kumar, R., Lang, A. E., & Lozano, A. M. (2000). An investigation of the effects of subthalamic nucleus stimulation on acoustic measures of voice. *Mov Disord, 15*(6), 1132-1138. doi:10.1002/1531-8257(200011)15:6<1132::aid-mds1011>3.0.co;2-o

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: their impact on measures of respiration, phonation, and articulation. *J Speech Lang Hear Res, 41*(5), 1003-1018. doi:10.1044/jslhr.4105.1003

Dromey, C., Ramig, L. O., & Johnson, A. B. (1995). Phonatory and articulatory changes associated with increased vocal intensity in Parkinson disease: a case study. *J Speech Hear Res, 38*(4), 751-764. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7474969

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. *J Commun Disord, 56*, 103-112. doi:10.1016/j.jcomdis.2015.06.012

Eichert, N., Papp, D., Mars, R. B., & Watkins, K. E. (2020). Mapping Human Laryngeal Motor Cortex during Vocalization. *Cereb Cortex, 30*(12), 6254-6269. doi:10.1093/cercor/bhaa182

Enderby, P. (2013). Disorders of communication: dysarthria. In M. Aminoff, F. Boller, & D. F. Swaab (Eds.), *Handbook of Clinical Neurology* (Vol. 110, pp. 273-281).

endolith. (2009). spectral_centroid.py. Retrieved from https://gist.github.com/endolith/359724

Engel, A. K., Moll, C. K., Fried, I., & Ojemann, G. A. (2005). Invasive recordings from the human brain: clinical insights and beyond. *Nat Rev Neurosci, 6*(1), 35-47. doi:10.1038/nrn1585

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci, 5*(1), 16-25. doi:10.1016/s1364-6613(00)01568-0

Favre, J., Burchiel, K. J., Taha, J. M., & Hammerstad, J. (2000). Outcome of unilateral and bilateral pallidotomy for Parkinson's disease: patient assessment. *Neurosurgery, 46*(2), 344-353; discussion 353-345. doi:10.1097/00006123-200002000-00017

Finnegan, E. M., Luschei, E. S., & Hoffman, H. T. (2000). Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech. *J Speech Lang Hear Res, 43*(4), 934-950. doi:10.1044/jslhr.4304.934

Fischer, P., Lipski, W. J., Neumann, W. J., Turner, R. S., Fries, P., Brown, P., & Richardson, R. M. (2020). Movement-related coupling of human subthalamic nucleus spikes to cortical gamma. *Elife, 9*. doi:10.7554/eLife.51956

Fischer, P., Pogosyan, A., Herz, D. M., Cheeran, B., Green, A. L., Fitzgerald, J., . . . Tan, H. (2017). Subthalamic nucleus gamma activity increases not only during movement but also during movement inhibition. *Elife, 6*. doi:10.7554/eLife.23947

Fogelson, N., Pogosyan, A., Kuhn, A. A., Kupsch, A., van Bruggen, G., Speelman, H., . . . Brown, P. (2005). Reciprocal interactions between oscillatory activities of different frequencies in the subthalamic region of patients with Parkinson's disease. *Eur J Neurosci, 22*(1), 257-266. doi:10.1111/j.1460-9568.2005.04179.x

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron, 88*(1), 220-235. doi:10.1016/j.neuron.2015.09.034

Gammon, S. A., Smith, P. J., Daniloff, R. G., & Kim, C. W. (1971). Articulation and stress-juncture production under oral anesthetization and masking. *J Speech Hear Res, 14*(2), 271-282. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/5558079

Gates, J. (2020). *The Ability of Persons with Parkinson's Disease to Manipulate Vocal Intensity and Articulatory Precision in an Intra-Operative Setting.* (Master of Science). University of Pittsburgh,

Gentil, M., Chauvin, P., Pinto, S., Pollak, P., & Benabid, A. L. (2001). Effect of bilateral stimulation of the subthalamic nucleus on parkinsonian voice. *Brain Lang, 78*(2), 233-240. doi:10.1006/brln.2001.2466

Georgopoulos, A. P., DeLong, M. R., & Crutcher, M. D. (1983). Relations between parameters of step-tracking movements and single cell discharge in the globus pallidus and subthalamic nucleus of the behaving monkey. *J Neurosci, 3*(8), 1586-1598. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6875658

Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron, 56*(6), 1127-1134. doi:10.1016/j.neuron.2007.09.038

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci, 15*(4), 511-517. doi:10.1038/nn.3063

Golfinopoulos, E., Cai, S., Blood, A., Burns, J., Noordzij, J. P., & Guenther, F. H. (2015). *Resting and task-based functional neuroimaging of adductor spasmodic dysphonia*. Paper presented at the Proceedings of the 2015 Human Brain Mapping Meeting, Honolulu, Hawaii.

Gordon, M. (1999). The effect of stress and prosodic phrasing on duration, acoustic amplitude and air flow of nasals in Estonian. *UCLA Working Papers in Phonetics, 92*, 151-159.

Grabski, K., Lamalle, L., Vilain, C., Schwartz, J. L., Vallee, N., Tropres, I., . . . Sato, M. (2012). Functional MRI assessment of orofacial articulators: neural correlates of lip, jaw, larynx, and tongue movements. *Hum Brain Mapp, 33*(10), 2306-2321. doi:10.1002/hbm.21363

Groenewegen, H. J. (2003). The basal ganglia and motor control. *Neural Plast, 10*(1-2), 107-120. doi:10.1155/NP.2003.107

Grover, T., Georgiev, D., Kalliola, R., Mahlknecht, P., Zacharia, A., Candelario, J., . . . Tripoliti, E. (2019). Effect of Low versus High Frequency Subthalamic Deep Brain Stimulation on Speech Intelligibility and Verbal Fluency in Parkinson's Disease: A Double-Blind Study. *J Parkinsons Dis, 9*(1), 141-151. doi:10.3233/JPD-181368

Guenther, F. H. (2015). *Neural control of speech*. Cambridge, MA: The MIT Press.

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., & Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Exp Brain Res, 130*(2), 133-141. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10672466

Haken, H., Peper, C., Beek, P. J., & Daffertshofer, A. (1996). A model for phase transitions in human hand movements during multifrequency tapping. *Physica D, 90*, 179-196.

Hammer, M. J., Barlow, S. M., Lyons, K. E., & Pahwa, R. (2010). Subthalamic nucleus deep brain stimulation changes speech respiratory and laryngeal control in Parkinson's disease. *J Neurol, 257*(10), 1692-1702. doi:10.1007/s00415-010-5605-5

Hammer, M. J., Barlow, S. M., Lyons, K. E., & Pahwa, R. (2011). Subthalamic nucleus deep brain stimulation changes velopharyngeal control in Parkinson's disease. *J Commun Disord, 44*(1), 37-48. doi:10.1016/j.jcomdis.2010.07.001

Hatsopoulos, N. G., Xu, Q., & Amit, Y. (2007). Encoding of movement fragments in the motor cortex. *J Neurosci, 27*(19), 5105-5114. doi:10.1523/JNEUROSCI.3570-06.2007

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*(2), 301-307. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14741110

Hebb, A. O., Darvas, F., & Miller, K. J. (2012). Transient and state modulation of beta power in human subthalamic nucleus during speech production and finger movement. *Neuroscience, 202*, 218-233. doi:10.1016/j.neuroscience.2011.11.072

Heida, T., Marani, E., & Usunoff, K. G. (2008). The subthalamic nucleus part II: modelling and simulation of activity. *Adv Anat Embryol Cell Biol, 199*, 1-85,vii. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18727495

Heilman, K. M., Leon, S. A., & Rosenbek, J. C. (2004). Affective aprosodia from a medial frontal stroke. *Brain Lang, 89*(3), 411-416. doi:10.1016/j.bandl.2004.01.006

Heracleous, P., Even, J., Ishi, C., Miyashita, T., Hagita, N., Kondo, M., & Takanohara, K. (2012). *The role of the Lombard reflex in parkinson's disease.* Paper presented at the IEEE 12th International Conference on Bioinfronmatrics & Bioengineering.

Hickok, G., & Poeppel, D. (2015). Neural basis of speech perception. *Handb Clin Neurol, 129*, 149-160. doi:10.1016/B978-0-444-62630-1.00008-1

Hillenbrand, J. M., Clark, M. J., & Nearey, T. M. (2001). Effects of consonant environment on vowel formant patterns. *J Acoust Soc Am, 109*(2), 748-763. doi:10.1121/1.1337959

Hoang, K. B., Cassar, I. R., Grill, W. M., & Turner, D. A. (2017). Biomarkers and Stimulation Algorithms for Adaptive Brain Stimulation. *Front Neurosci, 11*, 564. doi:10.3389/fnins.2017.00564

Horn, A., & Kuhn, A. A. (2015). Lead-DBS: a toolbox for deep brain stimulation electrode localizations and visualizations. *Neuroimage, 107*, 127-135. doi:10.1016/j.neuroimage.2014.12.002

Hoyte, K. J., Brownell, H., & Wingfield, A. (2009). Components of speech prosody and their use in detection of syntactic structure by older adults. *Exp Aging Res, 35*(1), 129-151. doi:10.1080/03610730802565091

Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., & Johnson, K. (1999). Formants of children, women, and men: the effects of vocal intensity variation. *J Acoust Soc Am, 106*(3 Pt 1), 1532-1542. doi:10.1121/1.427150

Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A. L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *Elife, 4*, e06213. doi:10.7554/eLife.06213

Iredale, J. M., Rushby, J. A., McDonald, S., Dimoska-Di Marco, A., & Swift, J. (2013). Emotion in voice matters: neural correlates of emotional prosody perception. *Int J Psychophysiol, 89*(3), 483-490. doi:10.1016/j.ijpsycho.2013.06.025

Isshiki, N. (1964). Regulatory Mechanism of Voice Intensity Variation. *J Speech Hear Res, 7*, 17-29. doi:10.1044/jshr.0701.17

Jannetts, S., & Lowit, A. (2014). Cepstral analysis of hypokinetic and ataxic voices: correlations with perceptual and other acoustic measures. *J Voice, 28*(6), 673-680. doi:10.1016/j.jvoice.2014.01.013

Jorge, A., Dastolfo-Hromack, C., Lipski, W. J., Kratter, I. H., Smith, L. J., Gartner-Schmidt, J. L., & Richardson, R. M. (2020). Anterior Sensorimotor Subthalamic Nucleus Stimulation Is Associated With Improved Voice Function. *Neurosurgery, 87*(4), 788-795. doi:10.1093/neuros/nyaa024

Joundi, R. A., Brittain, J. S., Green, A. L., Aziz, T. Z., Brown, P., & Jenkinson, N. (2012). Oscillatory activity in the subthalamic nucleus during arm reaching in Parkinson's disease. *Exp Neurol, 236*(2), 319-326. doi:10.1016/j.expneurol.2012.05.013

Jurgens, U. (2002). Neural pathways underlying vocal control. *Neurosci Biobehav Rev, 26*(2), 235-258. doi:10.1016/s0149-7634(01)00068-9

Jurgens, U. (2009). The neural control of vocalization in mammals: a review. *J Voice, 23*(1), 1-10. doi:10.1016/j.jvoice.2007.07.005

Jurgens, U., & Zwirner, P. (1996). The role of the periaqueductal grey in limbic and neocortical vocal fold control. *Neuroreport, 7*(18), 2921-2923. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9116210

Jwair, S., Coulon, P., & Ruigrok, T. J. (2017). Disynaptic Subthalamic Input to the Posterior Cerebellum in Rat. *Front Neuroanat, 11*, 13. doi:10.3389/fnana.2017.00013

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. *J Speech Lang Hear Res, 57*(4), 1178-1190. doi:10.1044/2014_JSLHR-S-13-0010

Karlsson, F., Olofsson, K., Blomstedt, P., Linder, J., & van Doorn, J. (2013). Pitch variability in patients with Parkinson's disease: effects of deep brain stimulation of caudal zona incerta and subthalamic nucleus. *J Speech Lang Hear Res, 56*(1), 150-158. doi:10.1044/1092-4388(2012/11-0333)

Kawato, M., & Wolpert, D. (1998). Internal models for motor control. *Novartis Found Symp, 218*, 291-304; discussion 304-297. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9949827

Keating, P. (2003, December 7-10, 2003). *Phonetic encoding of prosodic structure.* Paper presented at the 6th International Seminar on Speech Production, Sydney.

Keating, P. (2006). Phonetic encoding of prosodic structure. In H. J. & T. M. (Eds.), *Speech Production Models, phonetic processes and techniques* (pp. 167-186): Psychology Press.

Keating, P., & Shattuck-Hufnagel, S. (2002). A Prosodic View of Word Form Encoding for Speech Production. *UCLA Working Papers in Phonetics, 101*, 112-156.

Kelso, J. A., & Tuller, B. (1984). Converging evidence in support of common dynamical principles for speech and movement coordination. *Am J Physiol, 246*.

Kemp, J. M., & Powell, T. P. (1971). The connexions of the striatum and globus pallidus: synthesis and speculation. *Philos Trans R Soc Lond B Biol Sci, 262*(845), 441-457. doi:10.1098/rstb.1971.0106

Kent, R., & Read, C. (2002). *The Acoustic Analysis of Speech*: Singular/Thomson Learning.

Klouda, G. V., Robin, D. A., Graff-Radford, N. R., & Cooper, W. E. (1988). The role of callosal connections in speech prosody. *Brain Lang, 35*(1), 154-171. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/3179701

Krishnan, A., & Parkinson, J. (2000). Human frequency-following response: representation of tonal sweeps. *Audiol Neurootol, 5*(6), 312-321. doi:10.1159/000013897

Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: representation of pitch contours in Chinese tones. *Hear Res, 189*(1-2), 1-12. doi:10.1016/S0378-5955(03)00402-7

Kuhn, A. A., Williams, D., Kupsch, A., Limousin, P., Hariz, M., Schneider, G. H., . . . Brown, P. (2004). Event-related beta desynchronization in human subthalamic nucleus correlates with motor performance. *Brain, 127*(Pt 4), 735-746. doi:10.1093/brain/awh106

Kunzle, H. (1976). Thalamic projections from the precentral motor cortex in Macaca fascicularis. *Brain Res, 105*(2), 253-267. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/816421

Lachaux, J. P., Fonlupt, P., Kahane, P., Minotti, L., Hoffmann, D., Bertrand, O., & Baciu, M. (2007). Relationship between task-related gamma oscillations and BOLD signal: new insights from combined fMRI and intracranial EEG. *Hum Brain Mapp, 28*(12), 1368-1375. doi:10.1002/hbm.20352

Lam, J., & Tjaden, K. (2013). Intelligibility of clear speech: effect of instruction. *J Speech Lang Hear Res, 56*(5), 1429-1440. doi:10.1044/1092-4388(2013/12-0335)

Lanciego, J. L., Luquin, N., & Obeso, J. A. (2012). Functional neuroanatomy of the basal ganglia. *Cold Spring Harb Perspect Med, 2*(12), a009621. doi:10.1101/cshperspect.a009621

Landau, A. N., Schreyer, H. M., van Pelt, S., & Fries, P. (2015). Distributed Attention Is Implemented through Theta-Rhythmic Gamma Modulation. *Curr Biol, 25*(17), 2332-2337. doi:10.1016/j.cub.2015.07.048

Larson, C. R., Sun, J., & Hain, T. C. (2007). Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F0 and amplitude control. *J Acoust Soc Am, 121*(5 Pt1), 2862-2872. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17550185

Lavoie, L. M. (2001). *Consonant strength : phonological patterns and phonetic manifestations*. New York: Garland Pub.

Lehiste, I., & Universität zu Köln. (1970). *Suprasegmentals*. Cambridge, Mass.,: M.I.T. Press.

Levelt, W. J. (1992). Accessing words in speech production: stages, processes and representations. *Cognition, 42*(1-3), 1-22. doi:10.1016/0010-0277(92)90038-j

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Brain and Behavioral Sciences, 22*(1), 1-38.

Levitt, J. S. (2014). A Case Study: The Effects of the "SPEAK OUT!" Voice Program for Parkinson's Disease. *International Journal of Applied Science and Technology, 4*(2).

Levy, E. S., Moya-Gale, G., Chang, Y. H. M., Freeman, K., Forrest, K., Brin, M. F., & Ramig, L. A. (2020). The effects of intensive speech treatment on intelligibility in Parkinson's disease: A randomised controlled trial. *EClinicalMedicine, 24*, 100429. doi:10.1016/j.eclinm.2020.100429

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36. doi:10.1016/0010-0277(85)90021-6

Liotti, M., Ramig, L. O., Vogel, D., New, P., Cook, C. I., Ingham, R. J., . . . Fox, P. T. (2003). Hypophonia in Parkinson's disease: neural correlates of voice treatment revealed by PET. *Neurology, 60*(3), 432-440. doi:10.1212/wnl.60.3.432

Lipski, W. J., Alhourani, A., Pirnia, T., Jones, P. W., Dastolfo-Hromack, C., Helou, L. B., . . . Richardson, R. M. (2018). Subthalamic Nucleus Neurons Differentially Encode Early and Late Aspects of Speech Production. *J Neurosci, 38*(24), 5620-5631. doi:10.1523/JNEUROSCI.3480-17.2018

Lirani-Silva, C., Mourao, L. F., & Gobbi, L. T. (2015). Dysarthria and Quality of Life in neurologically healthy elderly and patients with Parkinson's disease. *Codas, 27*(3), 248-254. doi:10.1590/2317-1782/20152014083

Liss, J. M., LeGendre, S., & Lotto, A. J. (2010). Discriminating dysarthria type from envelope modulation spectra. *J Speech Lang Hear Res, 53*(5), 1246-1255. doi:10.1044/1092-4388(2010/09-0121)

Lizarazu, M., Lallier, M., & Molinaro, N. (2019). Phase-amplitude coupling between theta and gamma oscillations adapts to speech rate. *Ann N Y Acad Sci, 1453*(1), 140-152. doi:10.1111/nyas.14099

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. *J Speech Hear Disord, 43*(1), 47-57. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/633872

Ludlow, C. L., Connor, N. P., & Bassich, C. J. (1987). Speech timing in Parkinson's and Huntington's disease. *Brain Lang, 32*(2), 195-214. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2961409

Lundgren, S., Saeys, T., Karlsson, F., Olofsson, K., Blomstedt, P., Linder, J., . . . van Doorn, J. (2011). Deep brain stimulation of caudal zona incerta and subthalamic nucleus in patients with Parkinson's disease: effects on voice intensity. *Parkinsons Dis, 2011*, 658956. doi:10.4061/2011/658956

Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *J Acoust Soc Am, 125*(6), 3962-3973. doi:10.1121/1.2990715

Marani, E., Heida, T., Lakke, E. A., & Usunoff, K. G. (2008). The subthalamic nucleus. Part I: development, cytology, topography and connections. *Adv Anat Embryol Cell Biol, 198*, 1-113, vii. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18727483

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. *Parkinsons Dis, 2015*, 382320. doi:10.1155/2015/382320

Mayer, J., Wildgruber, D., Reicker, A., Dogil, G., Ackermann, H., & Grodd, W. (2002). *Prosody production and perception: converging evidence from fMRI studies.* Paper presented at the International Symposium on Computer Architecture, Aix-en-Provence, France.

Mazzoni, P., Hristova, A., & Krakauer, J. W. (2007). Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J Neurosci, 27*(27), 7105-7116. doi:10.1523/JNEUROSCI.0264-07.2007

Meigh, K. M. (2017). A Novel Investigation of Generalized Motor Program Theory: Syllable Stress as a Motor-Class Variable. *J Speech Lang Hear Res, 60*(6S), 1685-1694. doi:10.1044/2017_JSLHR-S-16-0247

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. *J Neurol Neurosurg Psychiatry, 78*(11), 1188-1190. doi:10.1136/jnnp.2006.110171

Mink, J. W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol, 50*(4), 381-425. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9004351

Monakow, K. H., Akert, K., & Kunzle, H. (1978). Projections of the precentral motor cortex and other cortical areas of the frontal lobe to the subthalamic nucleus in the monkey. *Exp Brain Res, 33*(3-4), 395-403. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/83239

Mugler, E. M., Goldrick, M., & Slutzky, M. W. (2014). Cortical encoding of phonemic context during word production. *Annu Int Conf IEEE Eng Med Biol Soc, 2014*, 6790-6793. doi:10.1109/EMBC.2014.6945187

Murdoch, B. E. (2001). Subcortical brain mechanisms in speech and language. *Folia Phoniatr Logop, 53*(5), 233-251. doi:10.1159/000052679

Nakayama, T., Fujii, Y., Suzuki, K., Kanazawa, I., & Nakada, T. (2004). The primary motor area for voluntary diaphragmatic motion identified by high field fMRI. *J Neurol, 251*(6), 730-735. doi:10.1007/s00415-004-0413-4

Nambu, A., Takada, M., Inase, M., & Tokuno, H. (1996). Dual somatotopical representations in the primate subthalamic nucleus: evidence for ordered but reversed body-map transformations from the primary motor cortex and the supplementary motor area. *J Neurosci, 16*(8), 2671-2683. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8786443

Nambu, A., Tokuno, H., Inase, M., & Takada, M. (1997). Corticosubthalamic input zones from forelimb representations of the dorsal and ventral divisions of the premotor cortex in the macaque monkey: comparison with the input zones from the primary motor cortex and the supplementary motor area. *Neurosci Lett, 239*(1), 13-16. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9547160

Narayana, S., Fox, P. T., Zhang, W., Franklin, C., Robin, D. A., Vogel, D., & Ramig, L. O. (2010). Neural correlates of efficacy of voice therapy in Parkinson's disease identified by performance-correlation analysis. *Hum Brain Mapp, 31*(2), 222-236. doi:10.1002/hbm.20859

Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., & Malach, R. (2007). Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Curr Biol, 17*(15), 1275-1285. doi:10.1016/j.cub.2007.06.066

Nowak, M., Zich, C., & Stagg, C. J. (2018). Motor Cortical Gamma Oscillations: What Have We Learnt and Where Are We Headed? *Curr Behav Neurosci Rep, 5*(2), 136-142. doi:10.1007/s40473-018-0151-z

Obeso, J. A., Rodriguez-Oroz, M. C., Rodriguez, M., Arbizu, J., & Gimenez-Amaya, J. M. (2002). The basal ganglia and disorders of movement: pathophysiological mechanisms. *News Physiol Sci, 17*, 51-55. doi:10.1152/nips.01363.2001

Olthoff, A., Baudewig, J., Kruse, E., & Dechent, P. (2008). Cortical sensorimotor control in vocalization: a functional magnetic resonance imaging study. *Laryngoscope, 118*(11), 2091-2096. doi:10.1097/MLG.0b013e31817fd40f

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience, 2011*. doi:10.1155/2011/156869

Panigrahi, B., Martin, K. A., Li, Y., Graves, A. R., Vollmer, A., Olson, L., . . . Dudman, J. T. (2015). Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell, 162*(6), 1418-1430. doi:10.1016/j.cell.2015.08.014

Patel, R., Niziolek, C., Reilly, K., & Guenther, F. H. (2011). Prosodic adaptations to pitch perturbation in running speech. *J Speech Lang Hear Res, 54*(4), 1051-1059. doi:10.1044/1092-4388(2010/10-0162)

Patel, R., Reilly, K. J., Archibald, E., Cai, S., & Guenther, F. H. (2015). Responses to Intensity-Shifted Auditory Feedback During Running Speech. *J Speech Lang Hear Res, 58*(6), 1687-1694. doi:10.1044/2015_JSLHR-S-15-0164

Patel, S., Lodhavia, A., Frankford, S., Korzyukov, O., & Larson, C. R. (2016). Vocal and Neural Responses to Unexpected Changes in Voice Pitch Auditory Feedback During Register Transitions. *J Voice, 30*(6), 772 e733-772 e740. doi:10.1016/j.jvoice.2015.11.012

Patel, S., Oishi, K., Wright, A., Sutherland-Foggio, H., Saxena, S., Sheppard, S. M., & Hillis, A. E. (2018). Right Hemisphere Regions Critical for Expression of Emotion Through Prosody. *Front Neurol, 9*, 224. doi:10.3389/fneur.2018.00224

Paul, L. K., Van Lancker-Sidtis, D., Schieffer, B., Dietrich, R., & Brown, W. S. (2003). Communicative deficits in agenesis of the corpus callosum: nonliteral language and affective prosody. *Brain Lang, 85*(2), 313-324. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12735947

Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody: evidence from ERPs. *Brain Res, 1217*, 171-178. doi:10.1016/j.brainres.2008.04.032

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain, 60*, 389-440.

Penney, J. B., Jr., & Young, A. B. (1983). Speculations on the functional anatomy of basal ganglia disorders. *Annu Rev Neurosci, 6*, 73-94. doi:10.1146/annurev.ne.06.030183.000445

Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22*, 227-250.

Perkell, J. S. (2013). Five decades of research in speech motor control: what have we learned, and where should we go from here? *J Speech Lang Hear Res, 56*(6), S1857-1874. doi:10.1044/1092-4388(2013/12-0382)

Perkell, J. S., Matthies, M., Lane, H., Guenther, F. H., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22*, 227-250.

Perkell, J. S., Matthies, M., Svirsky, M. A., & Jordan, M. I. (1995a). Goal-Based Speech Motor Control - a Theoretical Framework and Some Preliminary Data. *Journal of Phonetics, 23*, 23-35.

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1995b). Goal-Based Speech Motor Control - a Theoretical Framework and Some Preliminary Data. *Journal of Phonetics, 23*(1-2), 23-35. doi:Doi 10.1016/S0095-4470(95)80030-1

Pfurtscheller, G., Neuper, C., & Kalcher, J. (1993). 40-Hz oscillations during motor behavior in man. *Neurosci Lett, 164*(1-2), 179-182. doi:10.1016/0304-3940(93)90886-p

Pichon, S., & Kell, C. A. (2013). Affective and sensorimotor components of emotional prosody generation. *J Neurosci, 33*(4), 1640-1650. doi:10.1523/JNEUROSCI.3530-12.2013

Pinto, S., Ferraye, M., Espesser, R., Fraix, V., Maillet, A., Guirchoum, J., . . . Debu, B. (2014). Stimulation of the pedunculopontine nucleus area in Parkinson's disease: effects on speech and intelligibility. *Brain, 137*(Pt 10), 2759-2772. doi:10.1093/brain/awu209

Pinto, S., Ghio, A., Teston, B., & Viallet, F. (2010). [Dysarthria across Parkinson's disease progression. Natural history of its components: dysphonia, dysprosody and dysarthria]. *Rev Neurol (Paris), 166*(10), 800-810. doi:10.1016/j.neurol.2010.07.005

Potter-Nerger, M., Reese, R., Steigerwald, F., Heiden, J. A., Herzog, J., Moll, C. K. E., . . . Volkmann, J. (2017). Movement-Related Activity of Human Subthalamic Neurons during a Reach-to-Grasp Task. *Front Hum Neurosci, 11*, 436. doi:10.3389/fnhum.2017.00436

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci U S A, 103*(20), 7865-7870. doi:10.1073/pnas.0509989103

Ramig, L., Halpern, A., Spielman, J., Fox, C., & Freeman, K. (2018). Speech treatment in Parkinson's disease: Randomized controlled trial (RCT). *Mov Disord, 33*(11), 1777-1791. doi:10.1002/mds.27460

Ramig, L. O. (1997). How effective is the Lee Silverman voice treatment? *ASHA, 39*(1), 34-35. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9008989

Ramig, L. O., Countryman, S., O'Brien, C., Hoehn, M., & Thompson, L. (1996). Intensive speech treatment for patients with Parkinson's disease: short-and long-term comparison of two techniques. *Neurology, 47*(6), 1496-1504. doi:10.1212/wnl.47.6.1496

Ramig, L. O., Fox, C., & Sapir, S. (2004). Parkinson's disease: speech and voice disorders and their treatment with the Lee Silverman Voice Treatment. *Semin Speech Lang, 25*(2), 169-180. doi:10.1055/s-2004-825653

Ramig, L. O., Fox, C., & Sapir, S. (2007). Speech disorders in Parkinson's disease and the effects of pharmacological, surgical and speech treatment with emphasis on Lee Silverman voice treatment (LSVT(R)). *Handb Clin Neurol, 83*, 385-399. doi:10.1016/S0072-9752(07)83017-X

Ramig, L. O., Fox, C., & Sapir, S. (2008). Speech treatment for Parkinson's disease. *Expert Rev Neurother, 8*(2), 297-309. doi:10.1586/14737175.8.2.297

Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'Brien, C., Hoehn, M., & Thompson, L. L. (2001). Intensive voice treatment (LSVT) for patients with Parkinson's disease: a 2 year follow up. *J Neurol Neurosurg Psychiatry, 71*(4), 493-498. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11561033

Ramig, L. O., Sapir, S., Fox, C., & Countryman, S. (2001). Changes in vocal loudness following intensive voice treatment (LSVT) in individuals with Parkinson's disease: a comparison with untreated patients and normal age-matched controls. *Mov Disord, 16*(1), 79-83. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11215597

Randazzo, M. J., Kondylis, E. D., Alhourani, A., Wozny, T. A., Lipski, W. J., Crammond, D. J., & Richardson, R. M. (2016). Three-dimensional localization of cortical electrodes in deep brain stimulation surgery from intraoperative fluoroscopy. *Neuroimage, 125*, 515-521. doi:10.1016/j.neuroimage.2015.10.076

Ray, N. J., Brittain, J. S., Holland, P., Joundi, R. A., Stein, J. F., Aziz, T. Z., & Jenkinson, N. (2012). The role of the subthalamic nucleus in response inhibition: evidence from local field potential recordings in the human subthalamic nucleus. *Neuroimage, 60*(1), 271-278. doi:10.1016/j.neuroimage.2011.12.035

Ray, S., Crone, N. E., Niebur, E., Franaszczuk, P. J., & Hsiao, S. S. (2008). Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J Neurosci, 28*(45), 11526-11536. doi:10.1523/JNEUROSCI.2848-08.2008

Rektorova, I., Barrett, J., Mikl, M., Rektor, I., & Paus, T. (2007). Functional abnormalities in the primary orofacial sensorimotor cortex during speech in Parkinson's disease. *Mov Disord, 22*(14), 2043-2051. doi:10.1002/mds.21548

Rektorova, I., Mikl, M., Barrett, J., Marecek, R., Rektor, I., & Paus, T. (2012). Functional neuroanatomy of vocalization in patients with Parkinson's disease. *J Neurol Sci, 313*(1-2), 7-12. doi:10.1016/j.jns.2011.10.020

Riecker, A., Ackermann, H., Wildgruber, D., Meyer, J., Dogil, G., Haider, H., & Grodd, W. (2000). Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional

magnetic resonance imaging (fMRI) study. *Brain Lang, 75*(2), 259-276. doi:10.1006/brln.2000.2356

Riecker, A., Wildgruber, D., Dogil, G., Grodd, W., & Ackermann, H. (2002). Hemispheric lateralization effects of rhythm implementation during syllable repetitions: an fMRI study. *Neuroimage, 16*(1), 169-176. doi:10.1006/nimg.2002.1068

Rimmele, J. M., Gross, J., Molholm, S., & Keitel, A. (2018). Editorial: Brain Oscillations in Human Communication. *Front Hum Neurosci, 12*, 39. doi:10.3389/fnhum.2018.00039

Rosa, M., Marceglia, S., Barbieri, S., & Priori, A. (2014). Local Field Potential and Deep Brain Stimulation. In *Encyclopedia of Computational Neuroscience* (pp. 1-20). Retrieved from https://link.springer.com/referenceworkentry/10.1007/978-1-4614-7320-6_547-1

Roussel, P., Godais, G. L., Bocquelet, F., Palma, M., Hongjie, J., Zhang, S., . . . Yvert, B. (2020). Observation and assessment of acoustic contamination of electrophysiological brain signals during speech production and sound perception. *J Neural Eng, 17*(5), 056028. doi:10.1088/1741-2552/abb25e

Rusz, J., Klempir, J., Baborova, E., Tykalova, T., Majerova, V., Cmejla, R., . . . Roth, J. (2013). Objective acoustic quantification of phonatory dysfunction in Huntington's disease. *PLoS One, 8*(6), e65881. doi:10.1371/journal.pone.0065881

Salari, E., Freudenburg, Z. V., Branco, M. P., Aarnoutse, E. J., Vansteensel, M. J., & Ramsey, N. F. (2019). Classification of Articulator Movements and Movement Direction from Sensorimotor Cortex Activity. *Sci Rep, 9*(1), 14165. doi:10.1038/s41598-019-50834-5

Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V., & Friederici, A. D. (2010). Prosody meets syntax: the role of the corpus callosum. *Brain, 133*(9), 2643-2655. doi:10.1093/brain/awq231

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 Lang, 87*(2), 253-258. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14585294

Sapir, S., Ramig, L. O., & Fox, C. M. (2011). Intensive voice treatment in Parkinson's disease: Lee Silverman Voice Treatment. *Expert Rev Neurother, 11*(6), 815-830. doi:10.1586/ern.11.43

Sapir, S., Ramig, L. O., Hoyt, P., Countryman, S., O'Brien, C., & Hoehn, M. (2002). Speech loudness and quality 12 months after intensive voice treatment (LSVT) for Parkinson's disease: a comparison with an alternative speech treatment. *Folia Phoniatr Logop, 54*(6), 296-303. doi:10.1159/000066148

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant centralization ratio: a proposal for a new acoustic measure of dysarthric speech. *J Speech Lang Hear Res, 53*(1), 114-125. doi:10.1044/1092-4388(2009/08-0184)

Sapir, S., Spielman, J. L., Ramig, L. O., Story, B. H., & Fox, C. (2007). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: acoustic and perceptual findings. *J Speech Lang Hear Res, 50*(4), 899-912. doi:10.1044/1092-4388(2007/064)

Schalling, E., Johansson, K., & Hartelius, L. (2017). Speech and Communication Changes Reported by People with Parkinson's Disease. *Folia Phoniatr Logop, 69*(3), 131-141. doi:10.1159/000479927

Scharine, A., & McBeath, M. (2019). Natural Regularity of Correlated Acoustic Frequency and Intensity in Music and Speech: Auditory Scene Analysis Mechanisms Account for Integrality of Pitch and Loudness. *Auditory Perception & Cognition, 1*(3-4), 205-228.

Schmidt, R. (1975). A Schema Theory of Discrete Motor Skill Learning. *Psychological Review, 82*(4).

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*(82), 225-260.

Schmidt, R. A. (2003). Motor schema theory after 27 years: reflections and implications for a new theory. *Res Q Exerc Sport, 74*(4), 366-375. doi:10.1080/02701367.2003.10609106

Schoffelen, J. M., Poort, J., Oostenveld, R., & Fries, P. (2011). Selective movement preparation is subserved by selective increases in corticomuscular gamma-band coherence. *J Neurosci, 31*(18), 6750-6758. doi:10.1523/JNEUROSCI.4882-10.2011

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci, 32*(1), 9-18. doi:10.1016/j.tins.2008.09.012

Sengupta, R., & Nasir, S. M. (2015). Redistribution of neural phase coherence reflects establishment of feedforward map in speech motor adaptation. *J Neurophysiol, 113*(7), 2471-2479. doi:10.1152/jn.00731.2014

Sengupta, R., & Nasir, S. M. (2016). The predictive roles of neural oscillations in speech motor adaptability. *J Neurophysiol, 115*(5), 2519-2528. doi:10.1152/jn.00043.2016

Sevald, C., Dell, G., & Cole, J. (1995). Syllable Structure in Speech Production: Are Syllables Chunks or Schemas? *Journal of Memory and Language, 34*(6), 807-820.

Shapiro, B. E., & Danly, M. (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain Lang, 25*(1), 19-36. doi:10.1016/0093-934x(85)90118-x

Spielman, J., Mahler, L., Halpern, A., Gilley, P., Klepitskaya, O., & Ramig, L. (2011). Intensive voice treatment (LSVT(R)LOUD) for Parkinson's disease following deep brain stimulation of the subthalamic nucleus. *J Commun Disord, 44*(6), 688-700. doi:10.1016/j.jcomdis.2011.05.003

Stalzman, E., & Munhall, K. (1989). A Dynamical Approach to Gesural Patterning in Speech Production. *Ecological Psychology*. doi:DOI: 10.1207/s15326969eco0104_2

Steinhauer, K. (2000). *The Relationship Among Voice Onset, Voice Quality and Fundamental Freuqency: A Dynamical Perspective.* (PhD). University of Pittsburgh,

Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: phonetic context effects. *J Acoust Soc Am, 122*(2), 1111-1129. doi:10.1121/1.2749716

Tadel, F., Bock, E., Niso, G., Mosher, J. C., Cousineau, M., Pantazis, D., . . . Baillet, S. (2019). MEG/EEG Group Analysis With Brainstorm. *Front Neurosci, 13*, 76. doi:10.3389/fnins.2019.00076

Takai, O., Brown, S., & Liotti, M. (2010). Representation of the speech effectors in the human motor cortex: somatotopy or overlap? *Brain Lang, 113*(1), 39-44. doi:10.1016/j.bandl.2010.01.008

Tan, H., Pogosyan, A., Anzak, A., Ashkan, K., Bogdanovic, M., Green, A. L., . . . Brown, P. (2013). Complementary roles of different oscillatory activities in the subthalamic nucleus in coding motor effort in Parkinsonism. *Exp Neurol, 248*, 187-195. doi:10.1016/j.expneurol.2013.06.010

Tan, H., Pogosyan, A., Anzak, A., Foltynie, T., Limousin, P., Zrinzo, L., . . . Brown, P. (2013). Frequency specific activity in subthalamic nucleus correlates with hand bradykinesia in Parkinson's disease. *Exp Neurol, 240*, 122-129. doi:10.1016/j.expneurol.2012.11.011

Tan, H., Pogosyan, A., Ashkan, K., Cheeran, B., FitzGerald, J. J., Green, A. L., . . . Brown, P. (2015). Subthalamic nucleus local field potential activity helps encode motor effort rather than force in parkinsonism. *J Neurosci, 35*(15), 5941-5949. doi:10.1523/JNEUROSCI.4609-14.2015

Tan, H., Pogosyan, A., Ashkan, K., Green, A. L., Aziz, T., Foltynie, T., . . . Brown, P. (2016). Decoding gripping force based on local field potentials recorded from subthalamic nucleus in humans. *Elife, 5*. doi:10.7554/eLife.19089

Tan, H., Zavala, B., Pogosyan, A., Ashkan, K., Zrinzo, L., Foltynie, T., . . . Brown, P. (2014). Human subthalamic nucleus in movement error detection and its evaluation during visuomotor adaptation. *J Neurosci, 34*(50), 16744-16754. doi:10.1523/JNEUROSCI.3414-14.2014

Tanaka, Y., Tsuboi, T., Watanabe, H., Kajita, Y., Fujimoto, Y., Ohdake, R., . . . Sobue, G. (2015). Voice features of Parkinson's disease patients with subthalamic nucleus deep brain stimulation. *J Neurol, 262*(5), 1173-1181. doi:10.1007/s00415-015-7681-z

Tanaka, Y., Tsuboi, T., Watanabe, H., Nakatsubo, D., Maesawa, S., Kato, S., . . . Sobue, G. (2020). Longitudinal Speech Change After Subthalamic Nucleus Deep Brain Stimulation in Parkinson's Disease Patients: A 2-Year Prospective Study. *J Parkinsons Dis, 10*(1), 131-140. doi:10.3233/JPD-191798

Terumitsu, M., Fujii, Y., Suzuki, K., Kwee, I. L., & Nakada, T. (2006). Human primary motor cortex shows hemispheric specialization for speech. *Neuroreport, 17*(11), 1091-1095. doi:10.1097/01.wnr.0000224778.97399.c4

Tilsen, S. (2009). Multitimescale dynamical interactions between speech rhythm and gesture. *Cogn Sci, 33*(5), 839-879. doi:10.1111/j.1551-6709.2009.01037.x

Tilsen, S. (2011). Effects of syllable stress on articulatory planning observed in a stop-signal experiment. *Journal of Phonetics, 39*, 642-649.

Tjaden, K. (2008). Speech and Swallowing in Parkinson's Disease. *Top Geriatr Rehabil, 24*(2), 115-126. doi:10.1097/01.TGR.0000318899.87690.44

Tjaden, K., Lam, J., & Wilding, G. (2013). Vowel acoustics in Parkinson's disease and multiple sclerosis: comparison of clear, loud, and slow speaking conditions. *J Speech Lang Hear Res, 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. *Am J Speech Lang Pathol, 26*(2S), 569-582. doi:10.1044/2017_AJSLP-16-0090

Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: acoustic and perceptual findings. *J Speech Lang Hear Res, 47*(4), 766-783. doi:10.1044/1092-4388(2004/058)

Tornqvist, A. L., Schalen, 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. *Mov Disord, 20*(4), 416-423. doi:10.1002/mds.20348

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Lang Cogn Process, 26*(7), 952-981. doi:10.1080/01690960903498424

Trail, M., Fox, C., Ramig, L. O., Sapir, S., Howard, J., & Lai, E. C. (2005). Speech treatment for Parkinson's disease. *NeuroRehabilitation, 20*(3), 205-221. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16340101

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. *Mov Disord, 29*(4), 532-538. doi:10.1002/mds.25816

Tripoliti, E., Strong, L., Hickey, F., Foltynie, T., Zrinzo, L., Candelario, J., . . . Limousin, P. (2011). Treatment of dysarthria following subthalamic nucleus deep brain stimulation for Parkinson's disease. *Mov Disord, 26*(13), 2434-2436. doi:10.1002/mds.23887

Tripoliti, E., Zrinzo, L., Martinez-Torres, I., Frost, E., Pinto, S., Foltynie, T., . . . Limousin, P. (2011). Effects of subthalamic stimulation on speech of consecutive patients with Parkinson disease. *Neurology, 76*(1), 80-86. doi:10.1212/WNL.0b013e318203e7d0

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. *Mov Disord, 23*(16), 2377-2383. doi:10.1002/mds.22296

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage, 16*(3 Pt 1), 765-780. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12169260

Tzagarakis, C., West, S., & Pellizzer, G. (2015). Brain oscillatory activity during motor preparation: effect of directional uncertainty on beta, but not alpha, frequency band. *Front Neurosci, 9*, 246. doi:10.3389/fnins.2015.00246

Van Lancker, D. (1980). Cerebral Lateralization of Pitch Cues in the Linguistic Signal. *Paper in Linguistics, 13*(2).

Van Lancker Sidtis, D., Pachana, N., Cummings, J. L., & Sidtis, J. J. (2006). Dysprosodic speech following basal ganglia insult: toward a conceptual framework for the study of the cerebral representation of prosody. *Brain Lang, 97*(2), 135-153. doi:10.1016/j.bandl.2005.09.001

Van Lieshout, P. (2004). Dynamical Systems Theory and it Application in Speech. In B. Massen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijin (Eds.), *Speech Motor Control in Normal and Disordered Speech*. Oxford: Oxford University Press.

van Wijk, B. C., Beek, P. J., & Daffertshofer, A. (2012). Neural synchrony within the motor system: what have we learned so far? *Front Hum Neurosci, 6*, 252. doi:10.3389/fnhum.2012.00252

Wang, E., Verhagen Metman, L., Bakay, R., Arzbaecher, J., & Bernard, B. (2003). The effect of unilateral electrostimulation of the subthalamic nucleus on respiratory/phonatory subsystems of speech production in Parkinson's disease--a preliminary report. *Clin Linguist Phon, 17*(4-5), 283-289. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12945603

Watson, P., & Montgomery, E. B., Jr. (2006). The relationship of neuronal activity within the sensori-motor region of the subthalamic nucleus to speech. *Brain Lang, 97*(2), 233-240. doi:10.1016/j.bandl.2005.11.004

Watson, P. J., & Hughes, D. (2006). The relationship of vocal loudness manipulation to prosodic F0 and durational variables in healthy adults. *J Speech Lang Hear Res, 49*(3), 636-644. doi:10.1044/1092-4388(2006/040)

Watts, C. R. (2016). A retrospective study of long-term treatment outcomes for reduced vocal intensity in hypokinetic dysarthria. *BMC Ear Nose Throat Disord, 16*, 2. doi:10.1186/s12901-016-0022-8

Weed, E., & Fusaroli, R. (2020). Acoustic Measures of Prosody in Right-Hemisphere Damage: A Systematic Review and Meta-Analysis. *J Speech Lang Hear Res, 63*(6), 1762-1775. doi:10.1044/2020_JSLHR-19-00241

Weintraub, S., Mesulam, M. M., & Kramer, L. (1981). Disturbances in prosody. A right-hemisphere contribution to language. *Arch Neurol, 38*(12), 742-744. doi:10.1001/archneur.1981.00510120042004

Welch, G. (1985). A Schema Theory of How Children Learn to Sing in Tune. *Psychology of Music*.

Wenke, R. J., Cornwell, P., & Theodoros, D. G. (2010). Changes to articulation following LSVT(R) and traditional dysarthria therapy in non-progressive dysarthria. *Int J Speech Lang Pathol, 12*(3), 203-220. doi:10.3109/17549500903568468

Wessel, J. R., Ghahremani, A., Udupa, K., Saha, U., Kalia, S. K., Hodaie, M., . . . Chen, R. (2016). Stop-related subthalamic beta activity indexes global motor suppression in Parkinson's disease. *Mov Disord, 31*(12), 1846-1853. doi:10.1002/mds.26732

Wildgruber, D., Ackermann, H., & Grodd, W. (2001). Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. *Neuroimage, 13*(1), 101-109. doi:10.1006/nimg.2000.0672

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage, 24*(4), 1233-1241. doi:10.1016/j.neuroimage.2004.10.034

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science, 269*(5232), 1880-1882. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7569931

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends Cogn Sci, 2*(9), 338-347. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21227230

Yuchen, J., Xiao, C., & Qingfang, Z. (2020). Theta band (4-8 Hz) oscillations reflect syllables processing in Chinese spoken word production. *Acta Psychologica Sinica, 52*(10).

Zauber, S. E., & Huber, J. E. (2015). *The Effects of Dopaminergic Drugs and Deep Brain Stimulation on Speech in Parkinson's Disease*. Paper presented at the International Congress of Movement Disorder Society.

Zavala, B., Brittain, J. S., Jenkinson, N., Ashkan, K., Foltynie, T., Limousin, P., . . . Brown, P. (2013). Subthalamic nucleus local field potential activity during the Eriksen flanker task reveals a novel role for theta phase during conflict monitoring. *J Neurosci, 33*(37), 14758-14766. doi:10.1523/JNEUROSCI.1036-13.2013

Zavala, B., Tan, H., Little, S., Ashkan, K., Green, A. L., Aziz, T., . . . Brown, P. (2016). Decisions Made with Less Evidence Involve Higher Levels of Corticosubthalamic Nucleus Theta Band Synchrony. *J Cogn Neurosci, 28*(6), 811-825. doi:10.1162/jocn_a_00934

Zavala, B., Zaghloul, K., & Brown, P. (2015). The subthalamic nucleus, oscillations, and conflict. *Mov Disord, 30*(3), 328-338. doi:10.1002/mds.26072

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 Phoniatr Logop, 61*(4), 189-199. doi:10.1159/000227996